



E-DATA & RESEARCH

Jaargang 9 | nummer 2

Nieuwsbrief over data en onderzoek in de alfa- en gamma-wetenschappen.

E-data & Research verschijnt drie keer per jaar en wordt mogelijk gemaakt door: Centerdata, CLARIAH, DANS, Huygens ING, de Koninklijke Bibliotheek en het RIVM.

INHOUD

- 2 Verslagen van events in Gehoord en bijgewoond
- 3 Kees Aarts over nationale data-infrastructuur
- 4 Deze datasets zijn ook sinds kort beschikbaar
- 5 LASA verzamelt en deelt verouderingsdata



- 5 Historisch letterkundige Stronks online op zoek
- 6 Creative Camps ontwikkelen tools
- 6 RDNL biedt ook in 2015 cursus aan
- 7 Historisch-geografische data toegankelijk
- 8 Ewoud over wel willen maar niet kunnen



Scan deze QR code met een smartphone om de website van E-data te bezoeken.
www.edata.nl

RemBench laat online mogelijkheden kunsthistorische data zien

Alles over Rembrandt digitaal doorzoekbaar

“Het resultaat is mooier dan verwacht.” Suzan Verberne, project-leider van Rembench, vertelt enthousiast over de website met informatie over het leven van schilder Rembrandt van Rijn.

Ilya Nieuwland

Met Rembench kun je informatie over Rembrandt (1606-1669) zoeken in vier databases: twee van de Rijksdienst voor Kunsthistorische Documentatie (RKD), een met kunstwerken en een met biografische informatie over kunstenaars; RUQuest, de bibliotheekcatalogus van de Radboud Universiteit (RU), en RemDoc, een verzameling met primaire bronnen.

Alles achter één interface

Zowel Verberne als Rudie van Leeuwen (RU) waren vanaf het begin betrokken bij RemBench en RemDoc, het project dat eraan vooraf ging. Van Leeuwen vertelt: “Een tijd geleden is, onder meer op initiatief van Volker Manuth (RU), besloten om samen met Het Rembrandthuis alle documenten die te maken hadden met het leven en werk van Rembrandt te digitaliseren. De keuze viel daarbij op eLaborate van Huygens ING, een omgeving voor het ontwikkelen van digitale edities. Remdoc was het resultaat van die inspanning. Vervolgens hebben we alle literaire bronnen en kunstwerken rond Rembrandt erbij betrokken. Daarvoor werkten we samen met Huygens ING, het RKD en onze Universiteitsbibliotheek.



Op dit zelfportret is Rembrandt circa 23 jaar (1629) bron Rembrandt [Public domain], via Wikimedia Commons

RemBench werd gefinancierd door CLARIN-NL en het doel was om alle relevante informatie uit RemDoc en RUQuest met die van de RKD te combineren achter één interface.”

Mappen van informatie

De uitdaging lag vooral in het mappen van informatie uit de verschillende databases. Zoals bijvoorbeeld bij mensen met verschillende rollen. Van Leeuwen noemt Samuel van Hoogstraaten. “Die is zowel kunstenaar in Rembrandts atelier als auteur van een boek waarin Rembrandts werk wordt behandeld. De faceted search op de site laat beide rollen zien. Die facets werken intuïtief; ze helpen je je zoekactie af te bakenen. Zo kun je

bijvoorbeeld ook alleen kunstwerken opvragen die zich in Sint Petersburg bevinden. Of alle doop-, trouw- en begrafenisakten van Rembrandts familie.”

Nieuwe ideeën

Vanuit het Huygens ING waren Bram Buitendijk en Gijsjan Brouwer verantwoordelijk voor het ontwikkelen van de site: “Voor een front-end ontwikkelaar was het een interessant project omdat we onze faceted search library hebben kunnen inzetten. Deze had altijd hetzelfde uiterlijk, maar voor RemBench hebben we een externe user interface-ontwerper ingehuurd. De nieuwe feedback heeft tot goede ideeën geleid. De faceted search is nu flexibeler en de vormgeving kan aangepast worden aan de toepassing.”

Meer dan een demo-site

Brouwer geeft, evenals Van Leeuwen, wel toe dat er nog wat ruwe randjes op de site te vinden zijn. Maar zowel Verberne als Van Leeuwen zijn erg enthousiast over het eindresultaat. Uit een gebruikersstudie met studenten kunstgeschiedenis bleek dat gebruikers positief waren over het gebruiksgemak van RemBench, de uitgebreide zoekopties en de grote hoeveelheid informatie die achter de website zit. Verberne: “Het resultaat is mooier en nuttiger geworden dan we oorspronkelijk hadden gedacht. Wat we nu hebben, is een voorbeeldapplicatie waarmee we kunnen laten zien wat er inmiddels mogelijk is met kunsthistorische data. Het is veel meer dan alleen een demo-site geworden.”

<http://rembench.huygens.knaw.nl/>

Actuele, eenduidige en betrouwbare informatie

Data experts buigen zich over informatiestelsel voor de zorg

Er wordt hard gewerkt aan een duurzaam informatiestelsel voor de zorg, ook door data experts. Inez Joung, senior adviseur Informatie-Infrastructuur Zorg bij het RIVM, vertelt. Rutger Nugteren

Eind 2013 informeerde minister Schippers de Tweede Kamer over een informatiestelsel voor de zorg. Een stelsel met actuele, eenduidige en betrouwbare informatie, waarmee ontwikkelingen in de Nederlandse zorg te volgen zijn. Een jaar later werd een bestuurlijk informatie-beraad opgericht. In dit beraad zetten het ministerie van VWS en zorgpartijen uit het veld samen stappen om tot dit stelsel te komen. Eén van de betrokken partijen, het RIVM, heeft van het ministerie de opdracht gekregen om een commu-

nity van data experts op te richten. Inez voert deze opdracht uit. “Deze community wordt gevormd door (data)experts afkomstig van landelijke registraties in de zorg, zoals zorgverzekeraars, overheidsinstanties, beleidsonderzoek en de patiëntenfederatie NPCF.”

Inzicht in (zorg)datasets

De eerste bijeenkomst afgelopen december was een succes. De 26 aanwezigen besloten dat beter inzicht in de bestaande registraties en datasets wenselijk is. Inez vervolgt: “Zorggegevens.nl, een website met informatie over gegevensbronnen die betrekking hebben op de volksgezondheid en zorg, zal hiertoe worden uitgebreid met meer gedetailleerde informatie over (zorg)datasets. Verder kan de community een vraagbaak worden voor vragen over

cijfers met betrekking tot gezondheid, zorg en zorggebruik. Bovendien kan de community nagaan waar samengewerkt kan worden op het gebied van open data vanuit de verschillende registraties en datasets.”

Gunstig voor onderzoek

Hoewel het primaire doel van het duurzame informatiestelsel in de zorg en de community van data experts zich niet richt op wetenschappelijk onderzoek, zou ook het wetenschappelijk onderzoek profijt kunnen hebben van deze ontwikkelingen. “Zo zal het toevoegen van data dictionaries aan zorggegevens.nl ook onderzoekers beter informeren welke registratie of databestand het beste gebruikt kan worden voor de beantwoording van een bepaalde onderzoeksvraag.

Eenduidigheid van taal, van begrip, en dus ook de uitwisselbaarheid van gegevens zal vergroot worden. Ook het publiceren van meer open access-datasets zal de toegankelijkheid van data voor onderzoeksdoel-einden vergroten. Dit alles is goed voor de BV Nederland: met elkaar werken we aan één actueel, eenduidig en betrouwbaar informatiestelsel voor de zorg.”

zorggegevens.nl/kennisplatform

Volg ons op Twitter

Volg de stakeholders van E-data ook online:
@DANSKNAW
@centerdata
@ClariahInfra
@HuygensING
@KB_Nederland
@RIVM



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 Vytupil (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: datasilo's 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 eenvoudig 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 deponeerders 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.

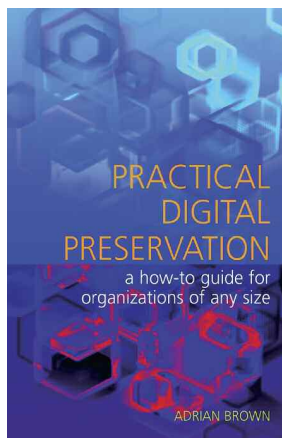
dans.knaw.nl

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 Dijkers. **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.

Kees Aarts, hoogleraar Politicologie Universiteit Twente:

‘Werken aan een infrastructuur binnen de sociale wetenschappen’

Kees Aarts, hoogleraar Politicologie Universiteit Twente, wil samenwerking en innovatie aanwakkeren in de Nederlandse maatschappij- en gedragswetenschappen door een nationale data-infrastructuur op te zetten. *Marika de Bruijne*

In de maatschappij- en gedragswetenschappen wordt op nationaal niveau overlegd over een nationale data-infrastructuur. Kees Aarts legt uit waarom: “De geesteswetenschappen hebben in de afgelopen jaren een enorme inhaalslag gemaakt als het gaat om innovatie in onderzoek. In het kielzog daarvan zijn er veel projecten gehonoreerd waardoor je ziet dat de geesteswetenschappen het op dit moment in een aantal opzichten beter doen dan de sociale wetenschappen. Daardoor vragen mensen in ons vakgebied zich nu af: wat gebeurt er als wij niets doen? Aan alle kanten doen zich tegelijk mogelijkheden voor om stappen te zetten. In de Wetenschapsvisie 2025 van het ministerie van Onderwijs, Cultuur en Wetenschap wordt een commissie voor onderzoeksfaciliteiten aangekondigd. De KNAW is bezig met een dergelijk plan. Het Europese Horizon 2020-programma kent veel initiatieven rond datamanagement en er zijn bewegingen zoals Open Access en ‘Science in transition’. Het gaat allemaal om de manier waarop we met gegevens omgaan.”

Wat zijn de plannen precies?

“Op dit moment is stap één ervoor zorgen dat wat we al hebben aan infrastructuur, niet verloren gaat. Een voorbeeld is het LISS panel. Dit panel van CentERdata heeft zich in de afgelopen zeven jaar bewezen als een belangrijke hoeksteen van Nederlands onderzoek. Dat geldt ook voor een aantal andere grootschalige onderzoeken zoals SHARE en European Social Survey. De volgende stap is om de vragende partijen, de instellingen die iets willen weten over de Nederlandse samenleving (zoals overheidsinstellingen en ministeries), bij elkaar aan tafel te krijgen, samen met universiteiten en onderzoeksinstellingen.” Deze partijen gaan vervolgens met elkaar in gesprek over mogelijke vormen van samenwerking en de voorwaarden die daaraan gesteld moeten worden. Het tijdspad: de eerste helft van dit jaar.

Hoe ziet de nieuwe data-infrastructuur eruit?

“Allereerst dient de infrastructuur een basis te bieden voor grootschalig survey-onderzoek in Nederland. Dit kan door een gemeenschappelijk kader te bieden waarin je onderzoek kunt afnemen. Door bijvoorbeeld een



Kees Aarts: “Aan alle kanten doen zich nu mogelijkheden voor” foto Bart van Vliet

samenwerking met het CBS zou je permanent voor goede steekproeven kunnen zorgen. Onderzoeken hoeft niet allemaal op dezelfde manier. Sommige surveys kun je face-to-face houden, sommige per post en sommige via internet. Dankzij gebruik van de beste kaders kun je de kwaliteit van de steekproeven echter verhogen en ervoor zorgen dat de gegevens die verzameld worden, elkaar kunnen verrijken.” Dit kan leiden tot kostenreductie. Maar Aarts benadrukt: “Het eerste doel van zo’n samenwerking is om het wetenschappelijke potentieel optimaal te benutten. Als het dan ook nog goedkoper kan dan is dat mooi meegenomen.”

Aarts vervolgt: “Door verschillende onderzoekers en onderzoeksgroepen met elkaar in contact te brengen, kunnen we innoveren in onze manier van onderzoeken. Elke onderzoeker heeft namelijk behalve iets innovatiefs ook iets conservatiefs in zich. We denken dat de manier waarop wij onze data verzamelen de enige juiste manier is, en dat iedereen die daaraan komt, het niet goed begrepen heeft. Hierdoor ontstaan idiomatische regels, zoals ‘deze studie moet altijd via face-to-face interviews worden afgenomen’. Als niet is onderzocht of dat inderdaad de beste methode is, dan is men niet innovatief bezig.

Door een toegankelijke mogelijkheid aan te bieden om met nieuwe werkwijzen te experimenteren, kun je onderzoekers van koudwaterrees afhelfen en kan de dialoog ontstaan.”

Hebben psychologen wel behoefte aan grootschalige surveyinstrumenten?

“Door meer met elkaar te praten en te kijken naar wat we kunnen delen over disciplines heen, kunnen de grenzen tussen dataverzamelingsstrategieën, tussen psychologie, sociologie en economie vervagen. Kijk bijvoorbeeld naar het LISS panel en waarom dat zo’n succes is geworden: omdat het een combinatie is van een langlopend representatief overzicht van opvattingen, houdingen en gedragingen van de Nederlandse bevolking. Echt een observatorium zou je kunnen zeggen. En tegelijk biedt het mogelijkheden om experimenten uit te voeren die psychologen eerder alleen met studenten konden uitvoeren. Ineens heb je geen scheiding meer tussen het experimentele paradigma en het observationele paradigma.”

“Er is een groep psychologen onder leiding van Gerben Westerhof die zich bezig houdt met narratieve psychologie. Met vraagstukken zoals: ‘hoe kunnen mensen door beter zelfinzicht zich beter voelen?’. Eén van de methoden die ze hebben bedacht, is om bewust te worden van je eigen levensverhaal, maar ook van dat van anderen. Dus de onderzoekers verzamelen levensverhalen. In het LISS panel is mensen gevraagd wat de drie meest bepa-

Kees Aarts

Kees Aarts is hoogleraar politicologie aan de Universiteit Twente en wetenschappelijk directeur van het Institute for Innovation and Governance Studies daar. Zijn belangstelling gaat uit naar democratie, verkiezingen, kiesgedrag en methodologie.

lende gebeurtenissen waren in hun leven tot nu toe. Dit waren open antwoorden die in verband zijn gebracht met persoonlijkheidskenmerken en andere achtergrondkenmerken. Door kwantitatieve en kwalitatieve dataverzamelingen zo te combineren kunnen vragen worden beantwoord die niet eerder zijn gesteld.”

Interdisciplinair werken, dat kunnen uw technische collega’s goed.

“De sociale wetenschappen worden gekenmerkt door een grote fragmentatie. In onze eigen subdisciplines denken wij als enigen te weten wat goed is. Ik heb op de Universiteit Twente veel contact met collega’s van technische faculteiten en zie hoe relatief gemakkelijk collega’s daar over disciplinaire grenzen heen kijken. Daar kunnen we nog veel van leren.

Zo gauw je samen een doel formuleert, zijn de onderlinge verschillen minder belangrijk. De vraag wordt dan hoe je het doel dichterbij kunt brengen. De gezamenlijke doelen mis ik nog zeer. Als sociale wetenschappers zouden we ons moeten richten op het afbreken van muurtjes tussen disciplines. Soms overdrijven we onze specialisaties. Pas als je bereid bent om te erkennen dat iemand uit de ene discipline, zeg psychologie, hele zinnige dingen kan zeggen over wat bijvoorbeeld een politicoloog met zijn kiezersonderzoek doet, dan heb je een gesprek.”

Voor meer informatie over de nieuwe infrastructuur, kunt u contact opnemen met Kees Aarts: c.w.a.m.aarts@utwente.nl.

LISS panel

Het LISS panel (Longitudinal Internet Studies for the Social sciences) is een online panel van Nederlandse huishoudens waarin universiteiten, onderzoeksinstituten, individuele wetenschappers en beleidsonderzoekers data kunnen verzamelen voor hun onderzoek. lissdata.nl

SHARE

De Survey of Health, Ageing and Retirement in Europe (SHARE) is een grootschalig Europees project dat sociale, economische en gezondheids-ervaringen van mensen van vijftig jaar en ouder onderzoekt. share-project.org

ESS

De European Social Survey (ESS) is een grootschalige studie naar de opvattingen van burgers over sociale, culturele en politieke kwesties binnen Europa. europeansocialsurvey.org

SINDS KORT BESCHIKBAAR

Dit overzicht toont databestanden die recent beschikbaar zijn gekomen bij CentERdata, Data Archiving and Networked Services en Huygens ING.

CentERdata



‘Eten’ was een thema binnen het onderzoek foto Shutterstock

• Leven in Nederland

Van 2011 tot en met 2012 is in het Immigrantenpanel van CentERdata een onderzoek afgenomen over hoe mensen het leven in Nederland ervaren. De data bestaan uit negen metingen met steeds een selectie uit verschillende thema's. De hoofdthema's zijn: eten, klimaat, mens, natuur, politiek, taal, vrijheid, werk en wonen. Het onderzoek is opgezet door Gamze Demirel, Suzan Elshout, Miquelle Marchand, Annette Scherpenzeel, Maarten Streefkerk en Corrie Vis. Door middel van de verschillende metingen wordt getracht een compleet beeld te krijgen van het leven in Nederland. Dit onderzoek werd gefinancierd door het MESS project, cross-culturele psychologie aan de Universiteit van Tilburg en Centraal Bureau voor de Statistiek (CBS). De data zijn beschikbaar via lissdata.nl.

Ook sinds kort beschikbaar:

Studies LISS panel

- Versluis, A.; Verkuil, B.; Brosschot, J.F., december 2013, Examining the relationship between effort-reward imbalance and implicit emotions in the workplace
- Terluin, B.; Brouwers, E.; Duijsens, I., juli - oktober - november 2013, Establishing norms for the Four-Dimensional Symptom Questionnaire (4DSQ) in the Dutch population, Wave 1-3
- Anderson, K.; Delsen, L.; Kantarcı, T., maart 2014, Stated Preference Analysis of Flexible Pension Plan Choices in the Netherlands
- Simons, A.M.W.; Groffen, D.A.I.; Bosma, H., februari 2013, Does stigmatisation 'explain' why low socioeconomic status is related to poor health?
- Fernee, H., september 2012 - september 2013, Time Use Study
- CentERdata, maart - april 2014, Family and Household - Wave 7
- CentERdata, april - mei 2014, Work and Schooling - Wave 7

Studies Immigrant panel

- Bekker, M.H.J.; Balkom, A.J.L.M. van; Ouwens, M.A.; Rutten, E.A.P., oktober - december 2012, Meaning of autonomy-connectedness in a multi-cultural (treatment) context - Part 1-3
- CentERdata, december 2013, Politics and Values - Wave 1
- Terluin, B.; Brouwers, E.; Duijsens, I., juli 2013, Establishing norms for the Four-Dimensional Symptom Questionnaire (4DSQ) in the Dutch population



Deze bestanden zijn kosteloos beschikbaar via www.lissdata.nl/dataarchive. Bezoek deze site of scan de QR-code.

DANS

• Nationaal Sportonderzoek

Het Nationaal Sportonderzoek (NSO) is de opvolger van de SportersMonitor, een onderzoek dat door het W.J.H. Mulier Instituut periodiek is uitgevoerd. De doelstelling van het NSO is het verkrijgen van inzicht in de sportparticipatie en de attitude ten aanzien van sportgerelateerde onderwerpen van de Nederlandse bevolking van 15 tot en met 80 jaar. De doelpopulatie bestaat derhalve uit alle Nederlanders in de leeftijd van 15 tot en met 80 jaar wonend in particuliere huishoudens. Eerdere metingen van de SportersMonitor hebben plaatsgevonden in 2008, 2010 en 2011. De onderzoeken van Drs. R. van den Dool (Mulier Instituut), Nationaal Sportonderzoek 2012 (2014-12-12) en 2013 (2014-09-23), persistent identifiers: [urn:nbn:nl:ui:13-owvg-2j](https://nbn-resolving.org/urn:nbn:nl:ui:13-owvg-2j) en [urn:nbn:nl:ui:13-4n0v-pv](https://nbn-resolving.org/urn:nbn:nl:ui:13-4n0v-pv) zijn beschikbaar in EASY.



Actieve ouderen tot 80 jaar betrokken in het onderzoek foto Shutterstock

Ook sinds kort beschikbaar:

- Verheijen, Dr. J.P.E. (2013-10-01), Balancing men, morals and money: Women's agency between HIV and security in a Malawi village. [urn:nbn:nl:ui:13-f11p-pn](https://nbn-resolving.org/urn:nbn:nl:ui:13-f11p-pn)
- SEO Economisch Onderzoek (2014-08-18), Effecten van een masteropleiding op leraren en hun omgeving. [urn:nbn:nl:ui:13-3hdg-dm](https://nbn-resolving.org/urn:nbn:nl:ui:13-3hdg-dm)
- Tolsma, Dr. J. (Radboud University Nijmegen); Kraaykamp, Prof. dr. G.L.M. (Radboud University Nijmegen); Graaf, Prof. dr. P.M. de (Universiteit Tilburg); Kalmijn, Prof. dr. M. (Universiteit van Amsterdam); Monden, dr. C.M. (Oxford University) (2014-11-03), Netherlands Longitudinal Lifecourse Study - NELLS Panel Wave 1 2009 and Wave 2 2013 - versie 1.1. [urn:nbn:nl:ui:13-5wyt-c6](https://nbn-resolving.org/urn:nbn:nl:ui:13-5wyt-c6)
- Sociaal Cultureel Planbureau - SCP; Centraal Bureau voor de Statistiek - CBS (2009), AVO-cultuur 2009, vervolgmodule 1: ICT-gebruik voor cultuur. [urn:nbn:nl:ui:13-4wnk-8w](https://nbn-resolving.org/urn:nbn:nl:ui:13-4wnk-8w)
- Sociaal Cultureel Planbureau - SCP; Centraal Bureau voor de Statistiek - CBS (2009), AVO-cultuur 2009, vervolgmodule 2: Kunstbeoefening. [urn:nbn:nl:ui:13-evsk-6f](https://nbn-resolving.org/urn:nbn:nl:ui:13-evsk-6f)
- Sociaal Cultureel Planbureau - SCP; Centraal Bureau voor de Statistiek - CBS (2009), AVO-cultuur 2009, vervolgmodule 3: Receptieve Cultuurparticipatie. [urn:nbn:nl:ui:13-2fik-w5](https://nbn-resolving.org/urn:nbn:nl:ui:13-2fik-w5)
- Centraal Bureau voor de Statistiek (2014-10), Gezondheidsenquête 2012. [urn:nbn:nl:ui:13-evl4-1m](https://nbn-resolving.org/urn:nbn:nl:ui:13-evl4-1m)
- Centraal Bureau voor de Statistiek (2014-10), Gezondheidsenquête 2013. [urn:nbn:nl:ui:13-3jen-dc](https://nbn-resolving.org/urn:nbn:nl:ui:13-3jen-dc)
- Haselager, dr. G.J.T. (Radboud University Nijmegen); Aken, prof. dr. M.A.G. van (Utrecht University) (2000), Personality and Family

Relationships; Gezin en Persoonlijkheid.

[urn:nbn:nl:ui:13-8mpw-ap](https://nbn-resolving.org/urn:nbn:nl:ui:13-8mpw-ap)

- As, MA S. van (2014-05), Opgraving Oss Horzak West 2013. Bewoningssporen uit de late ijzertijd en de late middeleeuwen. [urn:nbn:nl:ui:13-mhxl-jq](https://nbn-resolving.org/urn:nbn:nl:ui:13-mhxl-jq)

[urn:nbn:nl:ui:13-j14l-qk](https://nbn-resolving.org/urn:nbn:nl:ui:13-j14l-qk)

- Tump, drs M (BAAC bv) (2014-09), Leersum, Middelweggebied, deelgebied 1 en 2. Opgraving. [urn:nbn:nl:ui:13-j14l-qk](https://nbn-resolving.org/urn:nbn:nl:ui:13-j14l-qk)



Via easy.dans.knaw.nl zijn deze bestanden beschikbaar. Bezoek deze site of scan de QR code.

Niet meer beschikbaar is de website edna.nl. Alle informatie over het e-depot voor de Nederlandse archeologie (EDNA) is opgenomen op de nieuwe DANS-website dans.knaw.nl. Bezoekers van edna.nl worden automatisch naar de juiste pagina op de DANS-website geleid.

Huygens ING

• Koninkrijk in wording online

Het Koninkrijk der Nederlanden werd geboren in onzekere en onoverzichtelijke tijden. Opstand en tegenstand, belegeringen en veldslagen, de vestiging van de monarchie en de totstandkoming van de Grondwet, de vorming van een nieuw bestuur en leger, diplomatieke onderhandelingen en de verdubbeling van het grondgebied – dit alles gebeurde naast en na elkaar in een periode van amper twee jaar. De website ‘Koninkrijk in wording’ wil in deze gebeurtenissen duidelijkheid brengen en inzicht verschaffen. Het beginpunt op deze site is de historische inleiding, die een beknopt chronologisch overzicht van de jaren 1813-1815 biedt. Voor wie meer wil weten, zijn er eveneens een calendarium met de belangrijkste data, een overzicht van relevante publicaties, directe toegang tot de originele bronnen, afbeeldingen van gebeurtenissen en personen en korte levensbeschrijvingen van de hoofdrolspelers.

koninkrijk1813.huygens.knaw.nl



Op huygens.knaw.nl staan deze en andere beschikbaar gekomen publicaties. Bezoek deze site of scan de QR-code.



Willem I in koningsmantel door Joseph Paelinck, 1819 bron Wikimedia Commons

AGENDA

9 - 11 maart • San Diego

Research Data Alliance. 5th Plenary meeting

De RDA werkt aan open toegang tot onderzoeksdata. Tijdens deze bijeenkomst wordt de voortgang van resultaten van de werkgroepen gepresenteerd. rd-alliance.org

9 - 12 maart • Utrecht

Week van het Digitaal Erfgoed

Een week lang aandacht voor het Nederlands digitaal erfgoed en voor de noodzaak tot nauwe samenwerking. weekvanhetdigitaal erfgoed.nl

11 - 13 maart • San Diego

Preservation and Archiving Special Interest Group

Tijdens deze meeting worden toepassingen op het gebied van digitale archivering gepresenteerd. asis.org/pasig

11 en 18 maart • Amsterdam

Brainstormsessie Grootchalige Onderzoeksfaciliteiten

Onderzoekers kunnen tijdens deze KNAW-middag hun ideeën presenteren voor een grootchalige onderzoeksfaciliteit voor grensverleggend wetenschappelijk werk na 2025. knaw.nl

13 maart • Hilversum

CLARIAH kick-off

Clariah gaat de infrastructuur voor digitale geesteswetenschappen verbeteren. Op 13 maart vindt de aftrap plaats van dit project. clariah.nl/clariah-kick-off

24 maart • Amsterdam

InterScience: Conflict en oplossingen

De Jonge (KNAW-)Akademie organiseert een serie symposia waarbij alfa-, bèta- en gammawetenschappers én kunstenaars hun visie op een actueel thema belichten. dejongeakademie.nl

19 - 22 mei • Los Angeles

Archiving 2015

Tijdens deze conferentie staan digitale images centraal: conversie, analyse, behoud en beheer. Gastheer van de conferentie is het Getty institute. imaging.org/ist/conferences

22 mei • Den Haag

5e ESS-workshop

Deze door DANS en het nationaal coördinatieteam van het European Social Survey (ESS) georganiseerde workshop biedt onderzoekers de gelegenheid om informatie uit te wisselen over de vele mogelijkheden van de ESS-data. dans.knaw.nl

2 - 5 juni • Minneapolis

IASSIST 2015

Het thema van dit jaarlijkse IASSIST event is 'Bridging the Data Divide: data in the International Context', vooral bedoeld voor professionals die diensten leveren voor onderzoek en onderwijs in de sociale wetenschappen. iassist2015.pop.umn.edu

8 - 9 juni • Antwerpen

DHBenelux - Conference for Digital Humanities Research

Deze conferentie op het gebied van digitale geesteswetenschappen wordt voor de tweede keer georganiseerd, dit keer in Antwerpen. dhbenelux.org

Onderzoeksgegevens van 5.000 ouderen beschikbaar

LASA verzamelt, analyseert, deelt en koppelt verouderingsdata

Volgens het CBS telt ons land over vier jaar zeven miljoen vijftigplussers. Dat is de helft van de volwassen Nederlandse bevolking. De data van LASA beslaan gegevens van meer dan 5.000 ouderen, ook beschikbaar voor andere onderzoekers. *Heidi Berkhout*

LASA staat voor Longitudinal Aging Study Amsterdam. LASA vertegenwoordigt rijke gegevens van een cohort-sequentieel, multidisciplinair langlopend onderzoek naar veroudering. De dataset beslaat gegevens over het sociaal, emotioneel, cognitief en fysiek functioneren van meer dan 5.000 ouderen in longitudinaal perspectief. De data geeft inzicht in vragen als: ‘hoe zit het met de participatie-samenleving, doen ouderen meer of minder aan vrijwilligerswerk dan 20 jaar geleden?’, ‘we zijn bezorgd over obesitas, maar hoeveel ouderen zijn ondervoed?’ en ‘markeren geheugenklachten het begin van dementie?’. In 2014 werd de dataset nog genomineerd voor de Nederlandse Dataprijs humaniora en sociale wetenschappen.

Data uit zes landen

Data van LASA wordt onder andere gebruikt in het European Project on OsteoArthritis (EPOSA; eposa.org). Erik Timmermans, onderzoeker in opleiding bij LASA, vertelt: “Het EPOSA-onderzoek richt zich op de determi-



Oud zijn is hip! credits Hein de Kort

nanten en persoonlijke en sociaal-maatschappelijke consequenties van artrose voor ouderen in Europa. Het EPOSA-onderzoek is een observationele studie waarin data over oude

ren (65-85 jaar) van cohortstudies uit zes Europese landen (Nederland, Duitsland, Italië, Spanje, Zweden en het Verenigd Koninkrijk) zijn samengebracht.”

Koppeling van data

De koppeling van data van verschillende cohortstudies over de gezondheid en het functioneren van ouderen in de algemene populatie met data over omgevingskenmerken levert waardevolle kennis en mogelijkheden op voor onderzoekers, beleidsmakers, stadsplanners en verzorgers. Erik vervolgt: “Zo zijn binnen het EPOSA-project ook secundaire data verzameld om onderzoek te doen naar de invloed van de leefomgeving op ouderen met artrose. Er is bijvoorbeeld onderzoek gedaan naar de invloed van het weer op gewrichtspijn van ouderen met artrose. Hiervoor zijn gegevens van lokale weerstations over temperatuur, neerslag, luchtdruk, luchtvochtigheid en windsnelheid per dag gekoppeld aan de gegevens van zogenoemde pijnkalenders uit het onderzoek. Daarnaast wordt binnen EPOSA een studie uitgevoerd naar de invloed van objectieve buurtkenmerken op de fysieke activiteit van ouderen met artrose. Voor dit onderzoek wordt geodata van het Geoplaza-platform (Geoplaza-platform, Vrije Universiteit, Amsterdam) gekoppeld aan accelerometrie-data van ouderen met en zonder artrose.”

Gebruik maken van LASA-data

Een overzicht van beschikbare data staat op de LASA-website. Hier staan ook formulieren om deze data aan te vragen. Daarnaast is een selectie van LASA-data beschikbaar in het digitale archief van DANS.

Voor vragen over de LASA-data kan contact worden opgenomen met de onderzoekers via lasa@vumc.nl.

lasa-vu.nl

Historisch letterkundige en ‘digitale fellow’ Els Stronks:

‘Werken met tools een hele kunst’

Historisch letterkundige Els Stronks ging in de KB op zoek naar de Nederlandse jeugd in gedigitaliseerde vroegmoderne teksten.

Steven Claeysens

“Rond 1630 schreef Jacob Cats ‘kinderen zijn hinderen’. Teksten als deze hebben het ideaal van ouderlijke controle en het schrikbeeld van ontsporende jeugd in het collectieve geheugen van Nederlanders geprent.” Aan het woord is Els Stronks, de tweede ‘digitale fellow’ van de Koninklijke Bibliotheek (KB) en het Netherlands Institute for Advanced Studies in the Humanities and Social Sciences (NIAS). “Tijdens het fellowship vroeg ik me af of we met digitale middelen een andere dynamiek tussen jong en oud in historische teksten kunnen blootleggen en daarmee ook het jeugdige elan in het 16e-, 17e- en 18e-eeuwse Nederland kunnen belichten. Wat

dachten jongeren zélf te kunnen? En zagen ouderen de jeugd wel uitsluitend als een leeftijdsgroep die geleid, gebogen of zelfs geknot moest worden?”

Data, metadata en tools

“Ik heb geprobeerd alle Nederlandstalige vroegmoderne teksten die al gedigitaliseerd zijn, en alle beschikbare bibliografische en biografische metadata, te gebruiken. Veel van die data en metadata worden momenteel door Delpher en Nederlab samengevoegd, en dat helpt enorm. Ook qua tools maakte ik gebruik van wat er beschikbaar is: van de stilistische analysetool Stylo tot meer inhoudelijk gerichte topic modelers als Mallet, sentimentmeters als LIWC en concordantiemakers als AntConc.”

Kinderen een probleem

“Mijn eerste bevindingen zijn dat in onderzoek naar denkbeelden over



Els Stronks: “Kunnen we met digitale middelen een andere dynamiek in historische teksten blootleggen en daarmee ook situaties anders belichten?” foto Jos Uljee, KB

de jeugd in vroegmoderne teksten – net zoals dat nu gebeurt in media als het over de moderne jeugd gaat –

sterk gefocust wordt op kinderen als een probleem. Maar er werden – zeker door jonge auteurs zelf – ook veel positievere gedachten over de jeugd geformuleerd.”

Kloof vragen en tools

“Het onderzoek heeft me ook geleerd dat de beschikbare tools hele grove overzichten opleveren van kenmerken van een tekst die een computer kan tellen, maar dat het nog een hele kunst is om onze complexe vragen om te vormen tot deelvragen die aanhaken bij die specifieke tekstkenmerken. Als het goed is, is de KB straks een Ngram-viewer voor de teksten in de Digitale Bibliotheek voor de Nederlandse Letteren (DBNL) en een RDF-graph voor de Short-Title Catalogue, Netherlands (STCN) rijker: twee tools waar ook andere onderzoekers veel mee kunnen.”

tinyurl.com/elsstronks

KNAW faciliteert discussie Visie Wetenschap kabinet

Eind 2014 bood de KNAW met de website podium.knaw.nl onderzoekers een online podium voor discussie over de wetenschapsvisie van het kabinet. Per thema uit de Wetenschapsvisie 2025 werden onderzoekers uitgenodigd hun mening te formuleren en te reageren op de meningen van anderen. Het KNAW-podium was actief van 25 november tot en met 31 december 2014. Reageren kan niet meer, nalezen van de reacties wel. De top drie van meest besproken onderwerpen: de Nationale Wetenschapsagenda, open access en de aangekondigde herinrichting van NWO. De vervolgstap is dat de KNAW de reacties betreft bij de discussie over de Wetenschapsvisie 2025 in de Tweede Kamer en met het ministerie van Onderwijs, Cultuur en Wetenschap. Blijf op de hoogte van de Wetenschapsvisie 2025 via de website van de KNAW. (HB)

knaw.nl

Aan de slag met Linked Data tijdens Creative Camps

Nieuwe tools voor Europese data

In oktober vond het eerste Creative Camp plaats van Talk of Europe. Verschillende teams ontwikkelden toepassingen die gebruikmaken van de Linked Data uit het project. *Erica Renckens*

Het project Talk of Europe (ToE) is een vervolg op PoliMedia, waarin verslagen van Nederlandse politieke debatten in combinatie met krantenartikelen en radiofragmenten doorzoekbaar zijn gemaakt. De doelstellingen van ToE zijn vergelijkbaar, maar ambitieuzer: het publiceren van de verslagen van het Europees Parlement als Linked Data en het stimuleren van de ontwikkeling van toepassingen om deze data te gebruiken.

Internationale dataset

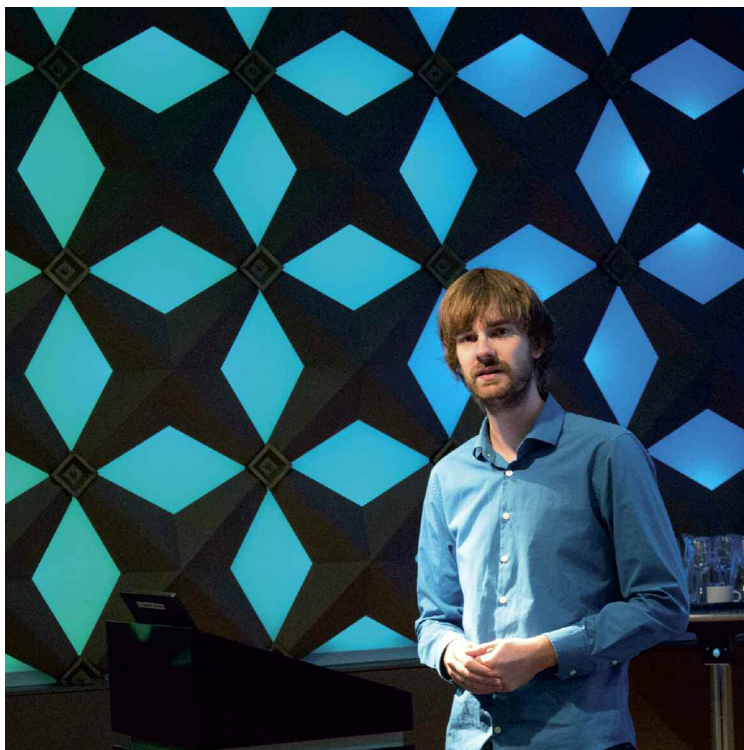
“De dataset met verslagen is al beschikbaar via een SPARQL-endpoint en de reacties zijn zeer positief”, aldus Martijn Kleppe, coördinator van het project namens de Erasmus Universiteit Rotterdam. “Deze internationale dataset is interessant voor onderzoekers in alle EU-landen.

Hij is beschikbaar in 22 talen, en kan zo gekoppeld worden aan datasets in net zoveel talen.”

Voor de ontwikkeling van toepassingen die gebruikmaken van deze dataset organiseert het projectteam drie Creative Camps. Gedurende vijf dagen werken teams uit verschillende landen aan tools voor onderzoekers uit de geestes- en sociale wetenschappen. Het eerste Creative Camp vond van 6 tot en met 10 oktober plaats in het Nederlands Instituut voor Beeld en Geluid.

Online beschikbaar

“Het Camp is zeer goed verlopen”, vertelt Max Kemman, organisator van het Camp namens de Erasmus Universiteit Rotterdam. De teams gingen ieder op verschillende manieren met de dataset aan de slag. “We zagen visualisaties om de da-



Organisator Max Kemman: “Met vijf teams was het Camp overzichtelijk. Dat gaf de mogelijkheid om met elkaar in discussie te gaan” foto Lindi Melse

taset inzichtelijk te maken, een koppeling met een nationaal parlement om politieke loopbanen te onderzoeken en een taalanalyse om de

dataset te verrijken of om meer inzicht te krijgen in hoe politici praten over onderwerpen.”

Sommige deelnemers hebben in-

middels hun tools online beschikbaar gesteld. Anderen willen ze eerst nog door ontwikkelen of willen ze inbouwen in bestaande producten.

Duidelijke meerwaarde

Het projectteam ziet een duidelijke meerwaarde in het organiseren van deze Creative Camps. “Als we de dataset met een persbericht online hadden gezet, was het wellicht wel opgepakt, maar had het waarschijnlijk lang geduurd voordat andere onderzoekers ermee aan de slag gingen”, aldus Kemman. “Daarnaast kregen we nu direct feedback op de data en hoe we deze nog beter kunnen communiceren. En ten slotte profiteren de deelnemers van elkaars kennis.”

Het volgende Camp vindt van 23 tot 27 maart plaats in het Meertens Instituut in Amsterdam.

Kleppe: “Dit keer formuleren we een aantal concrete thema’s, zodat de ontwikkelde tools verder aansluiten bij de behoeften van geestes- en sociale wetenschappers.”

talkofeurope.eu

Cursus ‘Essentials 4 Data Support’ ook in 2015 door dataprofessionals te volgen

RDNL ondersteunt de ondersteuner

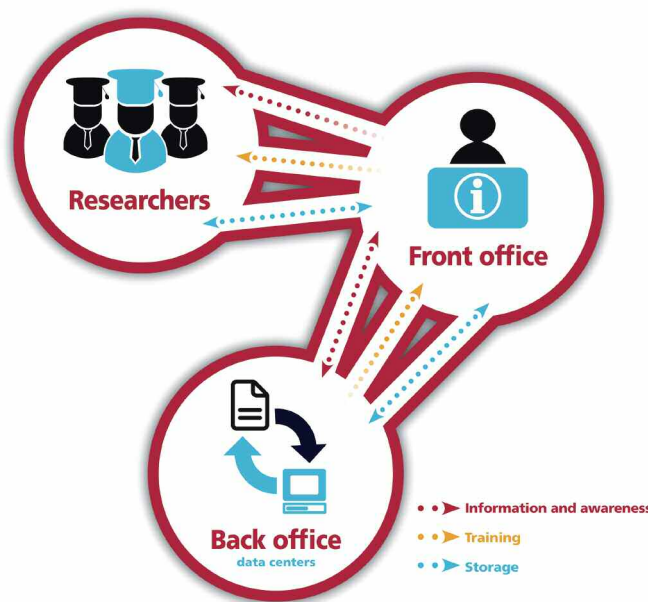
Voor dataprofessionals die onderzoekers beter willen begeleiden bij het duurzaam opslaan en delen van data biedt Research Data Netherlands een cursus aan.

Heidi Berkhout

In 2011 werd de training ‘Data Intelligence 4 Librarians’ ontwikkeld. Madeleine de Smaele, data steward bij 3TU.Datacentrum: “Met deze training boden we bibliotheekmedewerkers hulp bij het ondersteunen van ‘hun’ onderzoekers in allerlei aspecten van datamanagement. 2011 was ook het jaar van het ‘Witboek dataprofessionals in Nederland’ en de daarin beschreven trend heeft zich voortgezet. Inmiddels wordt de cursus, omgedoopt tot ‘Essentials 4 Data Support’, aangeboden door Research Data Netherlands (RDNL, zie kader, red.) en zijn we ruim zestig deelnemers verder.”

Datarapport centraal

De cursus richt zich op datasupporters: iedereen die onderzoekers wil helpen bij het opslaan, beheren, archiveren en delen van onderzoeksdata. Naast bibliotheekmedewerkers melden zich steeds vaker data librarians en IT-medewerkers voor de cursus aan, evenals data-managers uit onderzoeksgroepen.



Het frontoffice-backofficemodel van Research Data Netherlands: de datasupporters in de onderzoeksinstelling (het frontoffice) vormen een essentiële schakel tussen onderzoekers en het duurzame archief (het backoffice) credits RDNL

FOCUS

‘Twee contactdagen en een online-zelfstudie’

“De training versterkt de expertise van deze datasupporters in de front-offices”, vertelt Marjan Grootveld van DANS, die samen met Ellen

Verbakel (3TU.Datacentrum) de cursisten coacht. “De laatste jaren wordt van onderzoekers meer verwacht op het gebied van zorgvuldige omgang met onderzoeksgegevens. De frontoffices adviseren en begeleiden deze onderzoekers. Samen zorgen ze ervoor dat onderzoeksdata goed gedocumenteerd in een duurzaam archief worden ondergebracht, op de lange termijn toegankelijk en beschikbaar voor hergebruik. Onderwerpen die aan bod komen tijdens de cursus variëren van de planning van datama-

nagement tot datacitatie en licenties voor datahergebruik. In gesprekken tussen onderzoeksfinanciers en RDNL komen zulke onderwerpen overigens ook aan de orde, zodat we de cursisten actuele informatie kunnen bieden.”

De cursus combineert twee contactdagen met online-zelfstudie. Die vorm geeft deelnemers enerzijds ruimte om eigen accenten te leggen, afhankelijk van hun leerdoel. Anderzijds leren ze tijdens de bijeenkomsten van de andere deelnemers en de gastsprekers en breiden ze hun netwerk uit. Daarnaast is de cursus ook ‘online-only’ te volgen.

Reacties van cursisten

De deelnemers beoordelen de cursus positief. “Prima online-only-cursus die opvalt door structuur, opbouw en helderheid”, aldus een cursist die de online-only-variant van de cursus doorliep. Over de volledige cursus, met de twee groepsbijeenkomsten, zei een deelnemer het volgende: “Door de combinatie van online materiaal, het ontmoeten van collega’s van andere instellingen en de praktijkoefeningen heb ik veel geleerd over uiteenlopende aspecten van datamanagement.” Marcel Ras, programmamanager van de Nationale Coalitie Digitale Duurzaamheid (NCDD), is één van de gastsprekers: “Het goede nieuws

is dat veel dataprofessionals hun kennis op het gebied van digitale duurzaamheid willen vergroten. Zij zien ook dat hun samenwerking met onderzoekers steeds belangrijker wordt. Bijblijven in kennis en expertise op het gebied van duurzame toegang tot digitale informatie is dan wel noodzakelijk. De NCDD constateert echter dat er in Nederland helaas nog nauwelijks opleidingen en trainingen voor deze professionals zijn. Gelukkig voorziet de cursus ‘Essentials 4 Data Support’ nu in deze behoefte. Zo dragen we samen bij aan kennisontwikkeling en kennisdeling als belangrijke pijlers voor duurzaam behoud van wetenschappelijke data.”

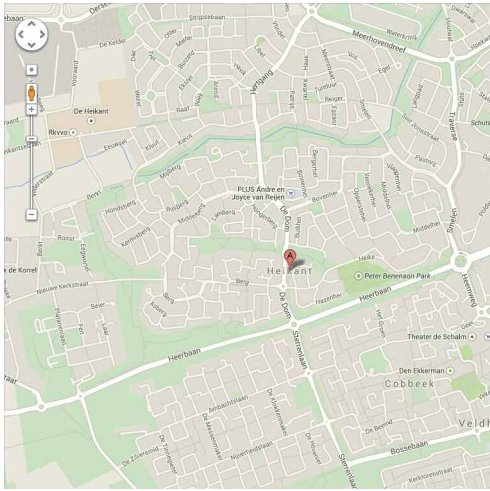
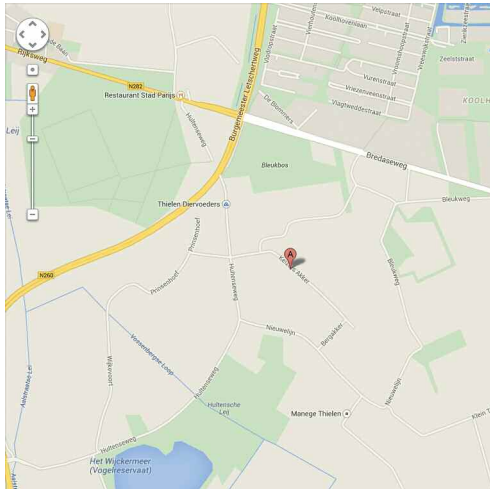
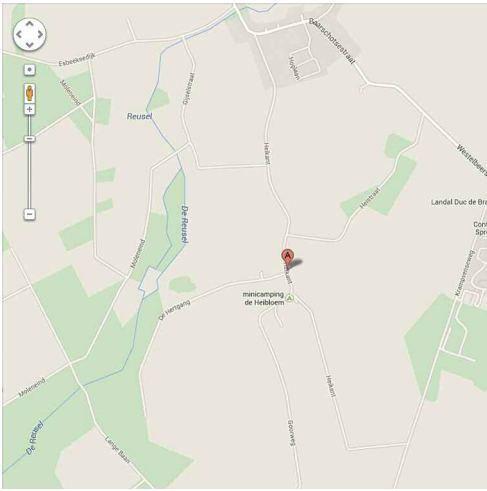
Ook in 2015 wordt de cursus twee keer aangeboden. Houd de website van RDNL in de gaten voor meer informatie.

researchdata.nl

RDNL

Research Data Netherlands (RDNL) is een samenwerkingsverband van 3TU.Datacentrum, DANS en SURFsara.

Met deze coalitie bundelen de drie partijen hun krachten op het gebied van duurzame data-archivering. Kijk op de website voor informatie over alle activiteiten: researchdata.nl.



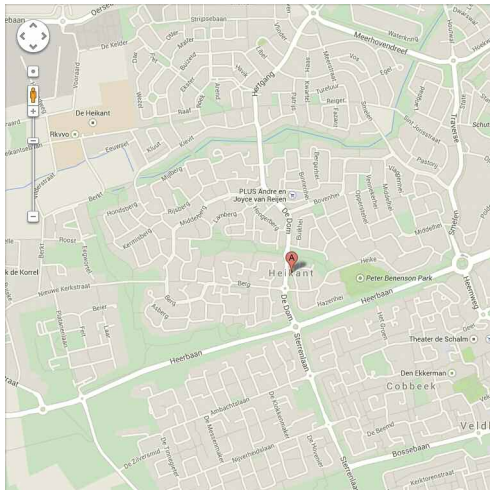
Kollektief Hic Sunt Leones versterkt digitale infrastructuur

Online hulp voor erfgoed

Geografische eenheden, zoals gemeenten, veranderen van naam en omvang. In 1921, bijvoorbeeld, gingen 10 buurgemeentes op in de stad Amsterdam. De mensen van Hic Sunt Leones, ICT-zelfstandigen met een belangstelling voor cultureel erfgoed, maken de bronnen waarop deze informatie is gebaseerd, toegankelijk.

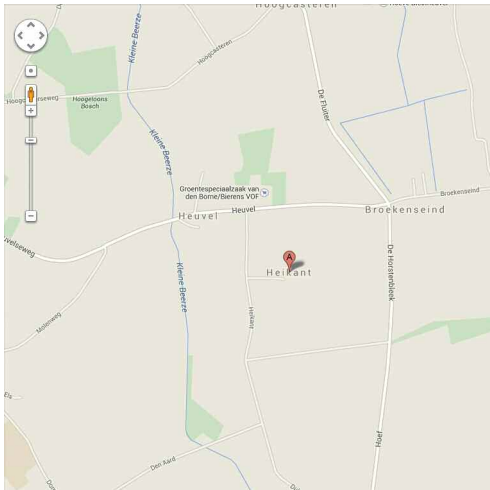
René van Horik

Zo biedt het kollektief met de website histopo.nl van meer dan 27.000 plaatsnaamvarianten de link naar de juiste plaats. De basis van dit systeem wordt gevormd door een aantal bronnen zoals het Repertorium van Nederlandse gemeenten (vanaf 1812) en de hieraan gerelateerde shape files die de coördinaten bevatten van gemeentegrenzen. Daarnaast zijn er andere bestanden opgenomen, bijvoorbeeld een overzicht van ‘verdwenen dorpen’ en plaatsnamen die opgenomen zijn in militieregisters. Bestanden met plaatsnamen, codes en coördinaten bestonden dus al langer, maar het kollektief



5 van de 24 ‘Heikant’-en in Nederland: als dorp, buurtschap of woonwijk, gevonden via histopo.nl bron googlemaps

heeft deze bestaande bronnen robuust, gebruiksvriendelijk en flexibel beschikbaar gesteld. Robuust, omdat elke geografische eenheid een duurzame codering heeft die gekoppeld kan worden aan externe bronnen. Gebruiksvriendelijk, omdat gemeentegediedenis.nl een prima gedocumenteerde website is. Met slechts een plaatsnaam of code (zelfs een postcode) is direct een overzicht beschikbaar van de geschiedenis van deze gemeente, alsmede een kaartje met daarop de grenzen van het gebied in een bepaald jaar. Flexibel, omdat informatie (zoals bijvoorbeeld alle plaatsen die eindigen op ‘beek’ door



middel van een API (Application Programming Interface) toegankelijk is voor ontwikkelaars. Bovendien ondersteunt de API verschillende uitvoerformaten, zodat optimale verdere verwerking mogelijk is.

Het bezoeken waard

De activiteiten van het kollektief maken historisch-geografische informatie toegankelijk voor een brede groep. Van de onderzoeker die data wil koppelen aan een geografische eenheid tot de liefhebber die genealogische gegevens in de tijd wil plaatsen, de website is een bezoek waard. hicsuntleones.nl

KORT

Miljoenen beelden voor de toekomst

Het project Beelden voor de Toekomst heeft de Nederlandse beeldgeschiedenis van de afgelopen honderd jaar gered van verval en teloorgang. De projectpartners hebben de afgelopen zeven jaar ruim 137.000 uur video, 22.000 uur film, 120.000 uur audio en bijna 3 miljoen foto's geconserveerd, gedigitaliseerd en toegankelijk gemaakt. Het audiovisueel materiaal is afkomstig uit de archieven van het Nederlands Instituut voor Beeld en Geluid, filminstituut EYE en het Nationaal Archief. Omdat de fysieke dragers van analoge opnames een beperkte levensduur hebben, dreigde het erfgoed verloren te gaan. Daarnaast was het overgrote deel van de collecties niet digitaal beschikbaar, zodat de mogelijkheden voor (her)gebruik beperkt bleven. Beelden voor de Toekomst heeft het audiovisueel materiaal nieuw leven ingeblazen en zo breed mogelijk beschikbaar gemaakt, zodat het toegankelijk blijft voor toekomstige generaties. Eind maart worden de resultaten van het project gepresenteerd tijdens een bijeenkomst in het Nederlands Instituut voor Beeld en Geluid. (ER) beeldenvoortetoekomst.nl



Archiefkasten met video's, film's, audio en foto's in Beeld en Geluid credits beeldenvoortetoekomst.nl

Privacyafwegingen internetgebruikers

Het verzamelen en gebruiken van persoonlijke gegevens door bedrijven en overheidsorganisaties roept steeds vaker reacties op bij burgers, die zich aangetast voelen in hun privacy. Anderzijds lijken mensen online relatief makkelijk persoonlijke informatie te delen met organisaties en via sociale media. Ook worden privacyvoorwaarden voor het gebruik van apps en online diensten vaak blind geaccepteerd. Uit onderzoek van CentERdata (in opdracht van ECP Platform voor de InformatieSamenleving) onder ruim 500 Nederlanders blijkt dat het krijgen van voordeel, zoals gratis verzending of een meer gebruiksvriendelijke app, mensen ertoe kan bewegen om persoonlijke informatie te delen, ook al vinden ze het gegevensgebruik eigenlijk onacceptabel.

Een erg groot voordeel kan echter argwaan opwekken (waarom heeft deze organisatie zoveel over voor mijn gegevens?) waardoor men juist minder bereid is om informatie te delen. Mensen die relatief makkelijk persoonlijke informatie vrijgeven, vinden privacy niet minder belangrijk, maar zij maken zich minder zorgen over hoe organisaties met de gegevens omgaan. De bereidheid om persoonlijke informatie te delen, lijkt dus niet zozeer voort te komen uit onverschilligheid of onwetendheid, maar uit een groter vertrouwen dat organisaties zorgvuldig met persoonlijke gegevens omgaan. Dit vertrouwen hangt sterk samen met opleiding en online ervaring; laagopgeleiden en mensen met weinig internetervaring zijn het meest bezorgd. (ME)

Vul s.v.p. deze antwoordkaart volledig in en stuur hem kosteloos op of ga naar edata.nl

BENT U TEVREDEN?

E-data&Research wordt per post (7500 ex) en digitaal (1900 ex) verspreid. Bent u tevreden, ontvangt u het blad liever (ook) op een andere manier of leest u tegenwoordig de digitale versie op edata.nl? Met deze antwoordkaart inventariseren we de wensen van onze lezers.

Ik ben tevreden. Dat is fijn, u hoeft de antwoordkaart NIET in te vullen.

Ik ontvang E-data&Research, maar voortaan liever anders, nl. :

- ☐ **digitaal** op het hieronderstaand e-mailadres, mijn postadres mag komen te vervallen.
- ☐ **per post** op het hieronderstaand adres, mijn e-mailadres mag komen te vervallen.
- ☐ **digitaal en per post** op de hieronderstaande adressen.
- ☐ **digitaal en/of per post** maar mijn gegevens kloppen niet meer, graag veranderen.
- ☐ **niet meer** middels deze antwoordkaart zeg ik mijn abonnement op.

Naam
Organisatie
Adres - oud
Adres - nieuw
E-mail - oud
E-mail - nieuw

Graag volledig invullen. Vragen? E-mail naar data@dans.knaw.nl. Bedankt voor uw reactie!

COLUMN

Tools: ik wil wel, maar waar moet ik precies beginnen?

In het vorige nummer van E-data schreef ik dat men in de e-humanities door de bomen het bos niet meer ziet. Te veel tools, te weinig samenwerking.

Er wordt wel degelijk geprobeerd om een geesteswetenschappelijke e-infrastructuur aan te leggen, luidt de samenvatting van een ingezonden brief in het huidige nummer. Zie de initiatieven en plannen voor de komende jaren van Virtual Language Observatory, CLAPOP, Bamboo Dirt en DASISH.

Vanzelfsprekend juich ik dergelijke initiatieven van harte toe, maar je kunt vier initiatieven om duidelijkheid te scheppen ook zien als meer van hetzelfde. Voor de gemiddelde gebruiker zal het er niet snel helderder op worden.

Onlangs gaf ik een gastcollege aan een paar van die gemiddelde gebruikers. In dit geval ging het om masterstudenten boekwetenschap. De vraag van de docent was of ik wilde laten zien hoe ik bepaalde tools gebruikte.

Ik liet onder meer zien hoe je met indexeringssoftware snel door grote datacollecties kunt zoeken – een van de weinige technieken die ik zelf redelijk beheers. Wildcards, jokertekens, woorden of namen in elkaars nabijheid zoeken, ik heb er dagelijks veel baat bij.

Vervolgens demonstreerde ik hoe je in Evernote heel makkelijk data kunt ordenen en van meta-data kunt voorzien.

Omdat dit alles bij veel master-

studenten tot glazige blikken leidde, stapte ik over op technieken om snel relevante bronnen op internet te vinden. Zoals via Google zoeken binnen een site, gebruikmaken van de url, zoeken naar bepaalde bestandstypen en filtertechnieken. ‘Had ik dat maar eerder in m’n studie geleerd, dat had me echt enorm veel tijd gescheeld’, verzuchtte een studente. Anderen beaamden dit.

Voor de duidelijkheid: voor zo-

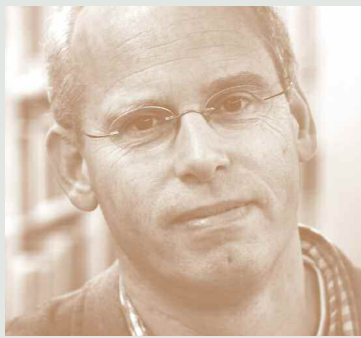


foto Leo van Velzen

ver ik dat kon beoordelen waren het allemaal leergierige en slimme mensen. Maar tegelijk was duidelijk dat ze niet erg vertrouwd waren met wat in mijn ogen inmiddels tot de basisvaardigheden van de e-humanities zou moeten behoren. Of eigenlijk: tot de basisvaardigheid van iedere student op een universiteit, hogeschool en middelbare school.

Natuurlijk heb ik gevraagd welke slimme tools ze gebruikten of kenden. Hun antwoorden, kort samengevat: van sommige tools hadden ze wel gehoord, maar ze gebruikten ze niet of nauwelijks want ze waren te moeilijk en het ontbrak aan

duidelijke uitleg.

Dat is ook mijn ervaring. Het is fijn dat er straks – over enkele jaren – websites zijn waarin je kunt zoeken naar honderden tools en honderdduizenden datacollecties. Maar ik zou het al erg fijn vinden als er nu een site zou bestaan met de tien beste en meest gebruikte tools voor bijvoorbeeld historisch en letterkundig onderzoek.

En niet alleen een lijstje met een korte beschrijving, maar iedere tool zou moeten worden voorzien van een heldere handleiding. Plus, nog belangrijker, een korte videocursus die je stap voor stap laat zien: 1. Hoe je de tool installeert, 2. wat je er in grote lijnen mee kunt, 3. hoe je een tool op maat kunt maken. En dan graag een paar datasets erbij zodat je zelf kunt ervaren hoe ongelooflijk nuttig zulk gereedschap kan zijn. Eis van de subsidiegever: de site moet minstens een keer per jaar grondig worden geactualiseerd.

Hoezeer dit ook voor de hand ligt, bij mijn weten ontbreekt het aan zo’n platform. Ik zou er zelf onmiddellijk gebruik van gaan maken, want er kan inmiddels veel meer dan ik weet. Maar net als voor die studenten geldt voor mij: ik wil wel, maar ik weet niet of nauwelijks waar ik moet beginnen.

Ewoud Sanders

Taalhistoricus en journalist. Sanders is vaste medewerker van onder meer NRC Handelsblad en Onze Taal.

Reactie op column Ewoud Sanders

In het woud van tools en data

Ewoud Sanders klaagt in zijn column ‘In de e-humanities ziet niemand door de bomen het bos meer’ (E-data, oktober 2014) dat er ‘nergens [...] een overzicht te vinden [is] van alle tools die zijn of worden ontwikkeld’. Hij heeft daar volkomen gelijk in.

Feitelijk is de situatie nog erger: er is ook geen overzicht van alle data die er zijn of die er gaan komen. Maar er zijn ook lichtpuntjes, zoals de Virtual Language Observatory (VLO), een poging van CLARIN (één van de initiatieven voor een geesteswetenschappelijke e-infrastructuur) om een dergelijk overzicht van alle data te maken. Dat is niet eenvoudig, omdat het onder andere vereist dat iedereen dezelfde standaarden gebruikt voor de beschrijving van de data. Er dient nog veel verbeterd te worden aan het overzicht en aan de manieren om in de data te zoeken, maar het is een goed begin. De portal bevat op dit moment ruim 800.000 metadatabestanden. VLO wordt systematisch verder uitgebreid en verbeterd met nieuwe zoekmogelijkheden.

Portal voor services

Daarnaast biedt CLARIN-NL met CLAPOP een online overzicht van inmiddels 40 software services die uit CLARIN-NL voortgekomen zijn,

waaronder TICCLops, MigMAP, OpenSONAR en War in Parliament. Daarnaast is er ook een overzicht met tot nu toe 25 datacollecties, zoals Discan en INTER-VIEWS. De portal stelt een gebruiker in staat naar deze data en services te zoeken op basis van vakgebied, taal, functionaliteit, en nog enkele criteria.

Internationaal

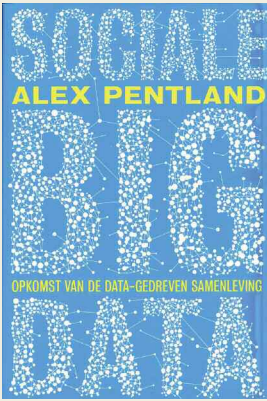
Ook internationaal wordt gewerkt aan brede catalogi voor tools en data in de geesteswetenschappen, zoals Bamboo Dirt. En ook in DASISH, een samenwerkingsverband van vijf grote onderzoeksinfrastructuren, worden dergelijke voorzieningen ontwikkeld. De zoekinterface en de kwaliteit van de metadata in zulke grootschalige overzichten moeten echter aan hoge eisen voldoen om nuttig te zijn voor onderzoekers uit de geesteswetenschappen. En dat is precies de uitdaging die vanuit CLARIN en CLARIAH de komende jaren wordt aangegaan. Ook al is er bij lange na nog geen volledig overzicht van ‘alle tools die zijn of worden ontwikkeld’, langzaamaan gaan we wel door de bomen het bos weer zien.

Jan Odijk - Directeur CLARIN-NL
Daan Broeder - Technisch directeur CLARIN-NL
portal.clarin.nl

GELEZEN

Sociale Big Data - Opkomst van de data-gedreven samenleving, Alex Pentland, 2014
Erica Renckens

De digitale revolutie biedt veel mogelijkheden voor de sociale wetenschappen, zo blijkt uit het boek Sociale Big Data van Alex Pentland, hoogleraar aan het MediaLab van het Massachusetts Institute of Technology (MIT), Verenigde Staten. Waar onderzoekers vroeger hun conclusies moesten trekken uit laagfrequente, veelal subjectieve metingen onder een select gezelschap, kan nu iedereen continu dienen als proefpersoon. De hele dag door laten we overal concrete sporen van ons gedrag achter, via sociale media, gps, telefoon- en betaalverkeer. Pentland beschrijft in zijn boek hoe hij in deze big data zoekt naar patronen om menselijk gedrag beter te kunnen beschrijven en beïnvloeden. ‘Sociale fysica’, zoals Pentland het noemt. Na een ietwat taaie introductie neemt Pentland de lezer in verschillende delen mee op een boeiende reis naar de hypotheti-



sche toekomst van organisaties, steden en samenlevingen, onder invloed van deze sociale fysica. Hij laat zien wat de uitkomsten van deze sociaal-fysische analyses kunnen betekenen voor de organisatie van bedrijven, wijken, steden en de gehele samenleving. De mogelijkheden klinken veelbelovend, maar er is één essentieel struikelblok: het privacyvraagstuk. Pentland pleit voor nieuwe wetgeving waarin de burger zelf zeggenschap heeft over zijn eigen data-profiel. Bepalen we straks zelf met wie we welke data delen? mavenpublishing.nl/boeken/sociale-big-data

Vul s.v.p. deze antwoordkaart volledig in en stuur hem kosteloos op of ga naar edata.nl

BENT U TEVREDEN?

Postzegel niet nodig

E-data&Research
Antwoordnummer 10275
1000 PA Amsterdam

