



E-DATA & RESEARCH

Jaargang 11 | nummer 3

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 Rijksmuseum.

INHOUD

2 Verslagen van events in Gehoord en bijgewoond

3 Digitaal taal materiaal en tools via INT

3 Vragen over data? Je kunt in Leiden terecht

4 Deze datasets zijn ook sinds kort beschikbaar



5 KNAW Cluster zet in op Digital Humanities

7 Online Media Suite biedt data en tools

8 Zelf datasets maken en delen met DataverseNL

8 Columnist Sanders stelt koppeling van data voor

E-data & Research wordt gratis toegezonden aan relaties van de stakeholders. Ook een uitgave ontvangen? Mail de redactie: edata@dans.knaw.nl.



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

Nieuwe afdeling Research Services verantwoordelijk voor alle collectiedata

Het Rijksmuseum als gereedschapskist

Sinds midden 2016 heeft het Rijksmuseum een nieuwe afdeling: Research Services. De speciale wetenschappelijke bibliotheek, collectiedocumentatie en IT en de Studie- en Leeszaal werken nu samen. Saskia Scheltjens

Op de westelijke zijwand van de monumentale bibliotheekhal in het Rijksmuseum staat in zwierige letters een spreuk: “Zijn Wetenschap en Kunst vaak met elkaar in Strijd, Hier is de Wetenschap der Kunst een zaal gewijd” (zie foto). Architect P.J.H. Cuypers liet deze spreuk tijdens de bouw in 1885 aanbrengen. De bezoeker wordt daarbij geheel volgens het negentiende-eeuwse wetenschapsideaal opgeroepen om zich bezig te houden met ‘de wetenschap van de kunst’. De nieuwe afdeling Research Services, gehuisvest rondom deze bibliotheekhal, neemt deze oproep ter harte en probeert daar een eenentwintigste-eeuwse invulling aan te geven. Met als afdelingsmissie ‘de ondersteuning van wetenschappelijk onderzoek in relatie tot de museumcollectie via open collectie-informatie en -data’.

Digitale uitdagingen

De (digitale) uitdagingen van Research Services zijn divers. Zo is het zaak om de bestaande digitale expertise op het vlak van linked open collectie(meta)data te consolideren en uit te breiden naar de bibliotheek- en onderzoeksdata en de collectiedocumentatie van het Rijksmuseum. De komende jaren zal daartoe met voorrang de technische infrastructuur en de personele bezetting worden verstevigd. Bouwen zonder een goed fundament heeft immers weinig zin. Er zal ook worden ingezet op een gefaseerde digitalisering van de Bijzondere Collecties van de Rijksmuseum Research Library en de analoge collectiedocumentatie. Parallel aan de digitalisering van de volledige objectcollectie tegen 2021, komen nu ook de bibliotheekcollectie en de documentatie in beeld. Deze wisselwerking kan verder nieuwe impulsen geven aan onderzoek naar de collectie.

Bundeling van krachten

Tot slot wordt er met diverse universitaire partners en onderzoeksinstituten gewerkt aan wetenschappelijk onderzoek naar (onderdelen van)



Computervisualisatie van de vergaderzaal van het Leidse dichtgenootschap ‘Kunst wordt door arbeid verkregen’ met centraal de reconstructie van het Panpoëticon Batavum. Het is een voorbeeld van interdisciplinair onderzoek waarbij fysieke objecten met nieuwe digitale technieken kunnen worden gevisualiseerd en gedeeld met een ruimer publiek.

credits Timothy De Paepe & Lieke van Deinsen, 2015

de collectie. Ook worden in de toekomst de logistieke collectiestromen en de fysieke toegang tot de afdeling verbeterd, inclusief de implementatie van scan-on-demand-diensten en een nieuwe ingang vanuit het atrium van het museum op de begane grond. Daarnaast zullen de bestaande publieke open-data-evenementen verder worden uitgebouwd.



De bekende neerlandicus, prof. dr. J. Verdam, schreef op verzoek van de architect van het Rijksmuseum, P.J.H. Cuypers, enkele programmatische spreuken die als wanddecoratie op de muren van de bibliotheek werden aangebracht. Hier de spreuk op de westelijke zijwand.

credits Rijksmuseum, 2013, CC01.0

Een recent voorbeeld van het interdisciplinaire onderzoek dat door de bundeling van krachten mogelijk wordt, is (een deel van) het promotie-onderzoek van voormalig Johan Huizinga Fellow (2015-2016) en huidige coördinator van de Studie- en Leeszaal van Research Services, Dr. Lieke van Deinsen. In haar proefschrift *Litteraire erf-laters. Canonvorming in tijden van culturele crisis, 1700-1750* keek zij, via verschillende invalshoeken, hoe de versnipperde collectie van een verdwenen achttiende-eeuws verzamelaarskabinet met daarin 350 portretten van Nederlandse schrijvers, het Panpoëticon Batavum, weer bij elkaar kon worden gebracht (zie afbeelding).

Onderzoek naar hoger niveau

Met de nieuwe afdeling wil het Rijksmuseum het digitaal, interdisciplinair en objectgericht onderzoek naar (Nederlandse) kunst en geschiedenis naar een hoger niveau tillen. Uniek ten opzichte van andere onderzoeksomgevingen is de combinatie van een internationaal opererend museum, een ambitieus onderzoeksinstituut rond conservering & restauratie (NICAS), objectgeoriënteerde inhoudelijke expertise en heel veel collectiedata.

rijksmuseum.nl/en/research

Saskia Scheltjens is hoofd van de afdeling Research Services van het Rijksmuseum.

Congres biedt staalkaart van methodologische en technische mogelijkheden

De grenzeloze gemeenschap van Digital Humanities Benelux

Begin juli komen ‘digitale’ geesteswetenschappers uit België, Nederland en Luxemburg voor de vierde keer bijeen. Na Den Haag (2014), Antwerpen (2015) en Luxemburg (2016) is Utrecht dit jaar het trefpunt. Steven Claeysens

De tijd dat de digitale geesteswetenschappen, in Nederland vooral bekend als Digital Humanities (DH), te klagen hadden over gebrek aan aandacht, ligt anno 2017 achter ons. De tijd dat ze nu en dan kreunen onder het gewicht van definities, meningen en de zonden Israëls, niet. Eén van

de oorzaken van die hardnekkige meerstemmigheid vormt tegelijk ook een grote aantrekkingskracht: Digital Humanities kent amper inhoudelijke grenzen en de beoefenaars zijn geschoold in de meest uiteenlopende vakgebieden.

Die grenzeloosheid maakt de digi-

tale humaniora vatbaar voor kritiek, maar zorgt ook voor een boeiende potpourri van sprekers en onderwerpen op de grote congressen. DHBenelux vormt daarop geen uitzondering. Max Kemman, promo-

Lees verder op pagina 2

GEHOORD & BIJGEWOOND

Onderzoekers helpen bij databeleid gemeenten

Patricia Prüfer

Wat gaat er veranderen door de steeds verdergaande digitalisering van de informatievoorziening en de samenleving? Hoe moeten overheidsinstellingen daarmee omgaan en op inspelen? Wat zou de meerwaarde kunnen zijn van beleidsvorming die meer op data gebaseerd is? En hoe moet men reageren op uitdagingen die verbonden zijn aan *data science*? Dit waren de thema's op het congres Digitale Agenda 2020 - Samen Organiseren, dat op 5 april jongstleden te Maarssen plaatsvond met 1.500 deelnemers uit meer dan driehonderd gemeenten. De Digitale Agenda 2020 (DA2020) is een initiatief van de Vereniging van Nederlandse Gemeenten (VNG) en het Kwaliteitsinstituut Nederlandse Gemeenten (KING) om gemeenten te ondersteunen bij en voor te bereiden op een sterk veranderende samenleving en verdergaande digitalisering. Om de *data science*-competentie van gemeenten te vergroten, kunnen gemeenten de hulp van onderzoekers goed gebruiken. Samen kan een datagedreven-beleid naar een hoger plan worden getrokken. In dit kader is vorig jaar de pilot Verkenning Data Lab voor gemeenten opgericht door VNG/KING en de Jheronimus Academy of Data Science (JADS). In een periode van vijf maanden gingen gemeenten onder wetenschappelijke begeleiding aan de slag met hun eigen data en vraagstukken. Ook werd de datavolwassenheid van gemeenten gescand. De uitkomsten van deze scan, *Data Science maturity van gemeenten*, werden, samen met



Onderdeel van de aanpak van VNG en KING is het verhaal Tegenwind. In deze thriller wordt duidelijk wat binnen de fictieve gemeente Turfbrug allemaal mogelijk is dankzij de ambitie, werkwijze, beweging en set van projecten van de Digitale VNG Agenda 2020. Download Tegenwind gratis via kinggemeenten.nl.
credits VNG/KING

aanbevelingen op het gebied van datavolwassenheid, gepresenteerd op DA2020. Daarnaast kwam het praktische gedeelte van de pilot aan bod. Gemeenten gingen samen met onderzoekers, data scientists en andere overheidsinstanties zoals de GGD zelf aan de slag

met data. De hackathon werd door deelnemers en toeschouwers als leuk, interessant en vruchtbaar ervaren. Op naar een vervolg in de KING Data Science Hub, de langetermijnsamenwerking die onlangs ook bij JADS in Den Bosch is gelanceerd.
da2020.nl/congres

Samenwerking, tools en data in 16 CLARIAH-pilots

Erica Renckens

De keynote speaker was door zijn rug gegaan en kon dus niet spreken tijdens de Toogdag van CLARIAH afgelopen maart bij het Internationaal Instituut voor Sociale Geschiedenis. Dit leek even een flink gat in het programma te slaan, maar bleek uiteindelijk vooral een zegening – al wenst iedereen de patiënt uiteraard beterschap. Het gaf de projectleiders van de zestien onlangs toegekende research pilots namelijk net wat extra tijd om hun plannen te presenteren. Elk jaar organiseert CLARIAH een Toogdag, waar het project een update geeft over de voortgang van de ontwikkeling van de digitale infrastructuur voor de geesteswetenschappen. Vrijdag 10 maart gaven ook de zestien onlangs toegekende research pilots een fraai beeld van de mogelijkheden die de data en tools in de nog volop in ontwikkeling zijnde infrastructuur bieden aan de verschillende vakgebieden. Ook maakten de presentaties duidelijk dat CLARIAH leidt tot veel nieuwe samenwerkingen, zowel binnen als tussen vakgebieden. Tot nu toe richtte CLARIAH zich op drie focusgebieden (taalkunde, sociaal-economische geschiedenis en mediastudies) met elk hun eigen

datatype (tekst, databases en audiovisueel materiaal). In de vervolgaanvraag, waaraan nu wordt gewerkt, komt hier nog een vierde focusgebied bij: inhoudelijke tekstanalyse. Na een demo-sessie van enkele bijna afgeronde projecten kregen de aanwezigen een update van de verschillende focusgebieden binnen CLARIAH en van de algemene technische ontwikkeling van de infrastructuur. Deze liggen mooi op schema en beloven nog veel goeds voor de komende jaren.
clariah.nl

LEARN-toolkit met voorbeelden op RDM-gebied voor onderzoekers

Ellen Leenarts

5 mei vond in Londen de eindconferentie van het LEARN project plaats. LEARN staat voor LEaders Activating Research Networks: Implementing the LERU Research Data Roadmap and Toolkit. Tijdens dit tweejarige project zijn ervaringen met Research Data Management (RDM) en open data verzameld van onderzoeksinstituten in Europa en Latijns-Amerika. Het resultaat is een toolkit, beschikbaar via de LEARN-website. Naast 23 praktijkvoorbeelden bevat de toolkit ook een concept RDM-beleid en een toelichting voor bestuurders en beslissers over het belang van RDM. Op dezelfde site is een webinar beschikbaar waarin de LEARN-toolkit kort wordt toegelicht. Bovendien staat op deze site een zogenoemde 'RDM-readiness'-test; een test waarmee op laagdrempelige wijze een indicatie wordt gegeven van RDM-thema's die binnen de eigen organisatie onderbelicht zijn.
learn-rdm.eu

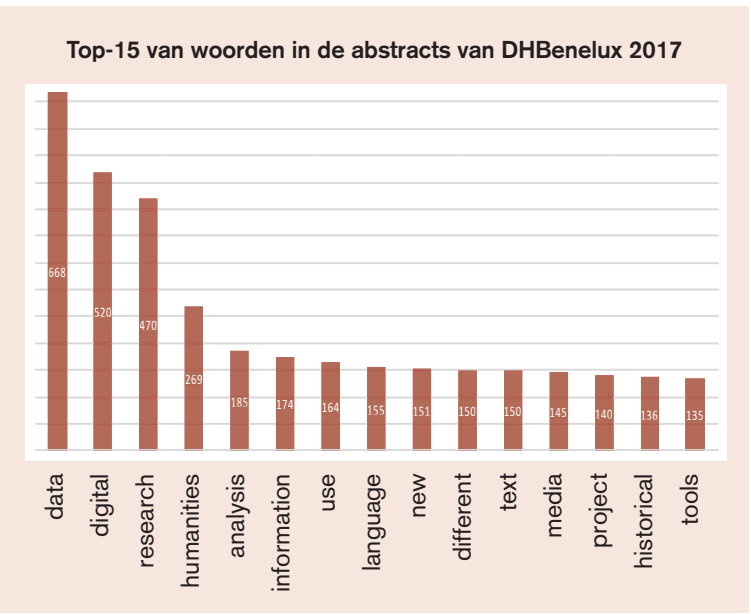
Digital Humanities Benelux-congres

Vervolg van pagina 1

vendus aan de Universiteit van Luxemburg, analyseerde de abstracts van deze editie en schreef erover op zijn blog maxkemman.nl. Hij constateert dat in 2017 het gemeenschappelijk woord bij uitstek 'data' is, één van de meest eenvoudige begrippen van deze tijd. Andere prominente woorden zijn 'digital', 'new', 'different', 'tools', 'project' en 'user', stuk voor stuk woorden met in de eerste plaats een methodologische signatuur. Kemman concludeert dat zeker de helft van de bijdragen door meerdere auteurs wordt geleverd, maar ook dat slechts een handvol onderzoekers ieder jaar een bijdrage levert. DHBenelux biedt met andere woorden een mix van elementen die per congres wisselt van samenstelling.

Doel van het congres

Het is eerder gesuggereerd: mischien is Digital Humanities wel op



Deze afbeelding toont hoe vaak 15 meest voorkomende woorden in de abstracts van DHBenelux 2017 worden genoemd. *credits Max Kemman*

de eerste plaats een gemeenschap. En bij de vorming van een wetenschappelijke gemeenschap – hoe digitaal ook – spelen conferenties een cruciale rol. Karina van Dalen-Oskam (Huygens ING en UvA), een

van de initiatiefnemers van de eerste editie: 'Ons doel was om onderzoekers in de Digital Humanities kennis te laten nemen van elkaars ervaringen en ze op de hoogte te stellen van de stand van zaken in

Nederland, Vlaanderen en daarbuiten op het terrein van inhoudelijke, methodologische en technische ontwikkelingen. Daarnaast wilden we de onderzoekers die nog niet betrokken zijn de mogelijkheid geven om een indruk te krijgen van de stand van zaken en om op zoek te gaan naar mogelijke samenwerkingspartners.'

Negen workshops

Het resultaat is deze vierde editie van DHBenelux, met paneldiscussies, demonstraties, posters en lezingen. Als opmaat voor het congres kan er dit jaar bovendien gekozen worden uit maar liefst negen workshops, variërend van het linken, annoteren of *minen* van data en media tot een sessie over Digital Humanities in de Arabische wereld. De keynotes worden verzorgd door Allan Marsden (Lancaster University) en Lora Aroyo (VU).
dhbenelux2017.eu

Rectificatie

In het vorige nummer van dit blad, rubriek Jong Talent, vertelt Ricarda Braukmann over datamanagement. Zij pleit onder andere voor het leggen van een link tussen data en proefschrift. Abusievelijk is vergeten te vermelden dat het rapport *Digitale objecten in digitale context. Eindrapport van het NCDD-project Digitale objecten in digitale context*, voorheen *Verrijkte publicaties* een van de aanleidingen voor Braukmann was om hier verder onderzoek naar te doen. In het onderzoek *Digitale objecten in digitale context* is de eerste aanbeveling om de verwijzing tussen proefschrift en data aan het begin van de dataketen te leggen (pagina 25). Wilt u meer weten? Het rapport is beschikbaar via de website van de NCDD, of neem contact op met Barbara Sierman, Digital Preservation Manager, Sector Innovation & Development bij de Koninklijke Bibliotheek.
ncdd.nl/site/wp-content/uploads/2015/08/Eindrapport-Digitale-objecten-in-digitale-context_DEF.pdf

OVERNEMEN ARTIKELEN

Wilt u een artikel uit dit blad overnemen? Dat mag altijd, maar vermeld wel de bron (E-data & Research) en de naam van de auteur van het artikel. Neem ook contact op met de hoofdredacteur (zie colofon) om door te geven waar artikelen geplaatst worden.

Instituut voor de Nederlandse Taal biedt data en tools

Onmisbaar voor de studie van het Nederlands van toen en nu

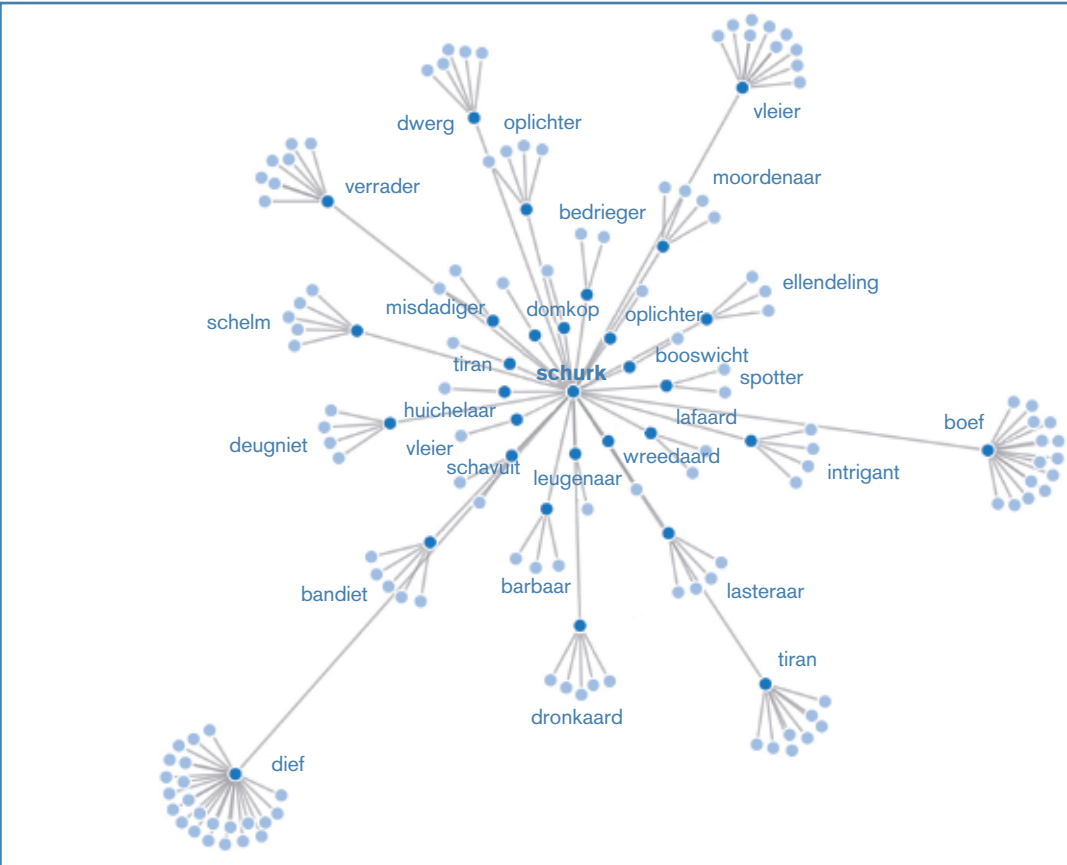
Wie het Nederlands bestudeert, kan er niet omheen: het digitale taalmateriaal en de onderzoekstools van het Instituut voor de Nederlandse Taal.

Wouter van Wingerden

Het Instituut voor de Nederlandse Taal (INT) is een voor het hele Nederlandse taalgebied breed toegankelijk wetenschappelijk instituut dat alle aspecten van het Nederlands bestudeert, er materiaal voor verzamelt, bewaart en bewerkt, en dat tools ontwikkelt voor taalonderzoek.

Onderzoek naar taal

Bij het INT kunnen taalwetenschappers onder meer terecht voor tekstcorpora en middelen om daarin onderzoek te doen. Een concreet voorbeeld van een actueel project is de bouw van een diachroon semantisch lexicon: DiaMaNT (zie afbeelding). Dit project is ingebed in het bredere Europese CLARIAH. De daaraan deelnemende CLARIN-centra – waaronder het INT – zijn voortrekkers op het gebied van geesteswetenschappelijk onderzoek. Ze hebben zich gecommitteerd aan gedeelde standaarden en hebben centrale diensten opgezet, die onderzoek goedkoper en efficiënter maken. Bovendien worden materialen, data en tools duurzaam bewaard en op een veilige manier beschikbaar gesteld. Ook werkt het instituut mee aan Digitisation.eu, dat het digitaliseren van historische teksten bevordert. Een van de lopende projecten is Nederlab, een webomgeving voor onderzoekers en studenten die zich bezighouden met historische ontwikkelingen in de Nederlandse taal, literatuur en cultuur. Al afgerond is het Taalportaal: de wetenschappelijke, Engelstalige grammatica van het Nederlands, Fries en Afrikaans. En momenteel wordt de Algemene Nederlandse Spraakkunst (ANS) uit 1997 grondig herzien.



Verborgen verbanden

Het project DiaMaNT (Diachroon seMANtisch lexicon van de Nederlandse Taal) is een hulpmiddel voor tekstontsluiting en bij het onderzoek naar begrippen door de eeuwen heen. Het lexicon legt automatisch relaties tussen woordvormen en betekenseenheden (concepten) en plaatst deze in de tijd. Deze afbeelding, afkomstig uit een demoversie, laat verbanden zien voor het woord 'schurk', op een manier die zowel mensen als computers begrijpen.
bron INT

ratuur en cultuur. Al afgerond is het Taalportaal: de wetenschappelijke, Engelstalige grammatica van het Nederlands, Fries en Afrikaans. En momenteel wordt de Algemene Nederlandse Spraakkunst (ANS) uit 1997 grondig herzien.

Rijke traditie

Verder heeft het INT een rijke traditie in het maken van woordenboeken. Voor het grote publiek is het immense *Woordenboek der Nederlandsche Taal*, voltooid in 1998, het bekendst. Het is samen met andere historische woordenboeken te raadplegen via gtb.ivdnt.org. Het hedendaagse Algemeen Nederlands

Woordenboek kun je zien groeien op anw.ivdnt.org. Maar ook die woordenboeken zijn nu ingebed in een internationale context: op dictionaryportal.eu zijn ze samen met wetenschappelijke woordenboeken van vele andere talen te raadplegen. Het INT maakt ook deel uit van het *European Network of e-Lexicography*, dat de zichtbaarheid, de toegankelijkheid en het gebruik van de Europese nationale woordenboeken tracht te vergroten.

Verzamelpplaats terminologie

Het INT verzamelt voortdurend neologismen (nieuwe woorden), waarmee bijvoorbeeld Woordenlijst.org wordt aangevuld (de uitge-

breide digitale versie van het Groene Boekje). Tot slot is het INT de verzamelplaats voor kennis over terminologie: woorden en begrippen uit allerlei soorten vaktaal in het Nederlands, gestandaardiseerd in Europese context.

Over al die activiteiten en voorzieningen is meer te lezen op de website. Daar kun je je ook inschrijven voor de nieuwsbrief van het instituut. En het laatste nieuws is er te lezen, bijvoorbeeld over het internationale congres eLex, dat in september plaatsvindt, over e-lexicografie in Leiden, de thuisstad van het INT.

ivdnt.org

Digitaal aankloppen in Leiden

In een paar jaar tijd openden in Leiden drie verschillende loketten voor onderzoekers met vragen over data en digitale onderzoeksmethoden. We zetten ze op een rij.
Heidi Berkhout

Leiden Centre of Data Science (LCDS)

Het LCDS, opgericht in 2014, is een netwerk van onderzoekers uit verschillende disciplines, die gebruik (willen) maken van innovatieve methodes voor het omgaan met data.

Onderzoekers met interesse voor innovatieve methoden en technieken voor het verzamelen en analyseren van grote datasets, kunnen terecht bij het LCDS. Het LCDS biedt expertise

en faciliteiten die nodig zijn voor dit type onderzoek. Naast het samenbrengen van onderzoekers uit verschillende wetenschapsgebieden werkt het LCDS samen met het bedrijfsleven en de overheid. Het LCDS organiseert regelmatig seminars voor onderzoekers en andere geïnteresseerden. Ook verzorgt het onderwijs op het gebied van Data Science.

Het LCDS is een initiatief van de Faculteit der Wiskunde en Natuurwetenschappen.

- DataScience LCDS
- DataScience LCDS
- universiteit leiden.nl/en/science/leiden-centre-of-datascience

Met dank aan: LCDS

Centre for Digital Scholarship (CDS)

De Universiteit Leiden heeft in juli 2016 het Centre for Digital Scholarship (CDS) opgericht om de impact van haar onderzoek en onderwijs te vergroten.

Onderzoekers kunnen bij het CDS terecht voor advies, training, workshops en het ontwikkelen van Proof of Concepts op een aantal expertisegebieden: Open Access, Datamanagement, Data Science en Digitale Duurzaamheid.

Het CDS draagt eveneens bij aan het verwezenlijken van de Open Science ambities van de Universiteit Leiden, zoals beschreven in het *Nationaal Plan Open Science*.

Het CDS is organisatorisch ondergebracht bij de Universitaire Bibliotheken Leiden.

- ubleiden
- ubleiden
- bibliotheek.universiteit leiden.nl/onderzoek-en-publiceren/centre-for-digital-scholarship

Met dank aan: Laurents Sesink

Leiden University Centre for Digital Humanities (LUCDH)

In april 2017 openden de deuren van het LUCDH, een verzamelpunt voor ondersteuning van onderzoek en onderwijs binnen de digitale geesteswetenschappen.

Geesteswetenschappers met onderzoeksplannen met een digitale component kunnen van het

LUCDH advies en ondersteuning krijgen op het terrein van: tekstanalyse, textmining, data-analyse en datavisualisering. Ook verzorgt het LUCDH het nieuwe BA-minorprogramma in Digital Humanities. Het centrum organiseert en sponsort activiteiten zoals: leesgroep, lunchtalks, workshops, summer schools en conferenties.

Het LUCDH is organisatorisch ondergebracht bij de Faculteit Geesteswetenschappen.

- leidendh
- LeidenDH
- universiteit leiden.nl/en/humanities/centre-for-digital-humanities

Met dank aan: Hilde De Weerd

SINDS KORT BESCHIKBAAR

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



foto Shutterstock

CentERdata

Hoe functioneren Nederlanders?
Hoe functioneren Nederlanders? Dit onderzoekt Martijn Oude Voshaar van de Universiteit Twente. In september 2015 heeft hij tachtig vragen voorgelegd aan een representatieve steekproef van de Nederlands sprekende bevolking. De vragen gaan over alledaagse activiteiten, van het dweilen van de keukenvloer of oppompen van de fietsband tot het oplopen van de trap. Met deze vragen meet Oude Voshaar beperkingen die men al dan niet ervaart bij het uitvoeren van algemene dagelijkse levensverrichtingen. Deze meting wordt gebruikt om de meeteigenschappen van de vragen te bepalen. Uiteindelijk hoopt de onderzoeker inzicht te krijgen in het functieniveau van de Nederlandse bevolking. De data van de studie *Computer Adaptive Test (CAT)* zijn beschikbaar via het LISS Data Archive. lissdata.nl

Ook sinds kort beschikbaar:
Studies LISS panel

- CentERdata, december 2016 - januari 2017, Politics and Values - Wave 9
- CentERdata, november 2016 - december 2016, Health - Wave 9
- CentERdata, augustus 2016 - september 2016, Religion and Ethnicity - Wave 9
- Jacobs, K., maart 2016 - april 2016, Election Survey Ukraine referendum, Measurement 1, Measurement 2 & Measurement 3

- Seuntjens, T.; Zeelenberg, M.; Breugelmans, S.; Ven, N. van de, september 2013, Dispositional Greed Scale



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

Huygens ING

Tijdschrift Gewina (1978-2007) online
Het *Tijdschrift voor de geschiedenis der Geneeskunde, Natuurwetenschappen, Wetkunde en Techniek*, later herdoopt tot *Gewina*, richtte zich speciaal op de wetenschapsgeschiedenis in Nederland en de voormalige koloniën. *Gewina* plaatste onder andere onderzoeksartikelen en artikelen over archiefvondsten, bijzondere collecties en nieuwe literatuur. *Gewina* werd van 1978 tot 1992 uitgegeven door Rodopi (Amsterdam) en verscheen tussen 1992 en 2007 bij Erasmus Publishing te Rotterdam. Sinds kort staan de uitgaves online. resources.huygens.knaw.nl/retroboeken/gewina

Ook sinds kort beschikbaar:
• *Scientiarum Historia: Tijdschrift voor de Geschiedenis van de Wetenschappen en de Geneeskunde* (1959-2006): resources.huygens.knaw.nl/retroboeken/scientiarumhistoria



Deze publicaties zijn beschikbaar via huygens.knaw.nl/tools-en-data/. Bezoek deze site of scan de QR-code.

DANS

Nieuwe informatie door Romeinse vondst
Prachtige restanten (zie foto) van Romeinse bronzen mensbeelden zijn gevonden in Naaldwijk. In de Romeinse tijd lag daar een grote nederzetting met connecties naar het nabijgelegen Forum Hadriani en het

kanaal van Corbulo. Kenmerkend is de grote hoeveelheid Romeins bronswerk die gedeeltelijk is omgesmolten, ook is er een kleine muntschat gevonden. Uit de vondsten lijkt verder naar voren te komen dat de plek ook in de 4e en 5e eeuw na Christus bewoond was, wat nieuwe informatie is. De dataset is door L.M.B. van der Feijst van ADC ArcheoProjecten via EASY beschikbaar gesteld.

<https://doi.org/10.17026/dans-24x-2nac>

Ook sinds kort beschikbaar:

De volgende datasets zijn beschikbaar via het online archiverings-systeem EASY van DANS:

- ResearchNed (2015): Studentenmonitor 2001-2015. DANS. <https://doi.org/10.17026/dans-zhz-dexe>
- Majid, Prof. Dr. A (Radboud University); Croijmans, MSc. I. (Radboud University) (2016): Human olfaction at the intersection of language, culture and biology. DANS. <https://doi.org/10.17026/dans-zke-2wgq>
- Kolipaka, S.S. (Leiden University) (2017): Diet and Prey Selection by Reintroduced Tigers in Multiple-use Forests. DANS. <https://doi.org/10.17026/dans-xwd-g6dh>
- Chilimampung, Dr C. (Chancellor College, University of Malawi) (2016): Acceptability and Feasibility of Early Infant Male Circumcision for HIV Prevention in Malawi. DANS. <https://doi.org/10.17026/dans-zxx-ye63>



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



De restanten van Romeinse bronzen mensbeelden, gevonden in Naaldwijk. credits ADC ArcheoProjecten

AGENDA

14 - 15 juni • Brighton

DPASSH 2017

Tijdens deze editie van 'Digital Preservation for Arts, Social Sciences and Humanities' staat de uitdaging 'Saving Everything' centraal. dpash.org

20 - 24 juni • Rotterdam

Scientific Visualization

Het eScience Center organiseert een congres met als thema 'From Archaeology to Astronomy'. esciencecenter.nl/event/scivis17

26 - 30 juni • Brisbane

International Conference on Open Repositories

Een conferentie georganiseerd voor en door de Open Repositories community. or2017.net

3 - 5 juli • Utrecht

Digital Humanities Benelux Conferentie

Deze jaarlijkse conferentie biedt een platform voor samenwerking tussen betrokkenen bij Digital Humanities projecten in België, Nederland en Luxemburg. dhbenelux2017.eu

8 - 11 augustus • Montréal

Digital Humanities

Het thema van deze Digital Humanities conferentie is 'Access'. dh2017.adho.org

10 - 13 september • Plymouth

Digital Research in the Humanities and Arts Conference (DRHA)

Deze editie van DRHA behandelt data in cross-disciplinaire context. drha2017.com

24 - 29 september • Warschau

Internationale EHRI-seminar

Tijdens dit seminar worden tools gepresenteerd nuttig voor onderzoek over de Tweede Wereldoorlog. ehri-project.eu

25 - 29 september • Kyoto

iPRES 2016

Deze editie van iPRES wordt georganiseerd door Kyoto University, Japan. ipres2017.jp

Privacy-groeimodel handig hulpmiddel voor universiteiten

Privacy: van de nood een deugd maken

Vanaf mei 2018 zijn organisaties die persoonsgegevens verwerken verplicht om aan een nieuwe Europese privacy-verordening te voldoen. Dit vraagt om goede voorbereiding. Het privacy-groeimodel helpt hierbij. Marlon Domingus

Het model kent vijf niveau's:

1. Initial, 2. Repeatable, 3. Defined, 4. Managed en 5. Optimised, beschouwd vanuit de volgende invalshoeken: universiteitsbreed, wetenschappelijke staf, juridische zaken, CIO en IT.

Privacy als kracht

In niveau 1 geldt dat iedereen heeft gehoord over privacy en datalekken

en hier vanuit de eigen werkzaamheden iets mee doet. In niveau 2 ontstaat een gemeenschappelijk beeld en wordt gezamenlijk naar oplossingen gezocht om privacy te borgen. In niveau 3 borgt de organisatie de bescherming van persoonsgegevens met standaardoplossingen. In niveau 4 biedt de organisatie naast standaardoplossingen ook maatwerk. In niveau 5 ziet de organisatie privacy niet langer als een externe dreiging voor de eigen taken, maar als een kracht.

Verskillende niveau's

Binnen één organisatie kunnen verschillende afdelingen zich op verschillende niveau's bevinden. Zo kan een IT-afdeling het gebruik van

Dropbox verbieden vanuit veiligheidsoogpunt, terwijl onderzoekers juist deze online clouddienst kunnen gebruiken voor het uitwisselen van data. Een Data Protection Officer

Algemene Verordening Gegevensbescherming

Er wordt ook gewerkt aan de Uitvoeringswet van de Algemene Verordening Gegevensbescherming (AVG). De precieze formulering van deze Uitvoeringswet, een nationale aanvulling op de AVG, is van groot belang voor wat straks mag met wetenschappelijke onderzoeksbestanden die persoonsdata bevatten. Meer hierover in een volgende *E-data & Research*.

buigt zich over een Privacy Impact Assessment en Juridische Zaken bekommert zich over de verplichting tot een register van verwerkingen van persoonsgegevens. Niveau 2 bundelt deze acties in een strategie om samen verantwoord met het gebruik van persoonsgegevens om te gaan.

25 mei 2018

Eén ding is duidelijk: op 25 mei 2018 moeten Nederlandse universiteiten zich niet alleen aan de regels houden, maar ook kunnen aantonen dat dit gebeurt. Het verantwoord uitvoeren en faciliteren van een big data onderzoeksproject waarin met persoonsgegevens wordt gewerkt, is vanaf niveau 3 geen werkelijke

uitdaging meer. Het is dan ook zaak om zo snel mogelijk in dit niveau terecht te komen. Gebruik het groeimodel hierbij, breng in kaart wat je zelf kunt doen en waar je (externe) hulp nodig hebt. Maak een planning en deel bevindingen over de voortgang. Voorbeelden, praktische handvatten, referentiekaarten en meer handige informatie staan op de website van het Landelijk Coördinatiepunt Research Data Management (LCRDM).

lcrdm.nl

Marlon Domingus is projectmanager researchdatamanagement aan de Erasmus Universiteit en ontwerper van het model.

Directeuren KNAW Humanities Cluster over bundeling van krachten:

‘Hoe meer data, hoe meer kennis: onze grenzen vervagen’

Sinds 1 oktober 2016 vormen het Huygens ING, het Meertens Instituut en het Internationaal Instituut voor Sociale Geschiedenis (IISG) samen het KNAW Humanities Cluster (HuC). Vooral op het gebied van ‘digital humanities’ hoopt het HuC een spilfunctie te gaan vervullen.

Mathilde Jansen

We zitten in de Commissariskamer van het Spinhuis: een historische plek in hartje Amsterdam. Een voormalig politiebureau, en ooit een huis waar veroordeelde vrouwen moesten spinnen en naaien. Sinds vorig jaar is het KNAW Humanities Cluster (HuC) hier gevestigd, althans twee van de drie instituten. Het IISG is vanwege de omvang van zijn archief- en bibliotheekcollectie op de Crispijnsweg in Amsterdam gebleven. Al ademt de Commissariskamer de sfeer van toen, de rest van het pand is strak ingericht met modern meubilair.

Die combinatie van traditioneel en modern vinden we ook terug in de onderzoeksmethodes van de drie instituten van het HuC. Modern staat hier voor digitaal, want digital humanities hebben de toekomst, daarover zijn de drie directeuren van het HuC het zonder meer eens. Dat is ook het punt waarop ze elkaar gevonden hebben en waarom ze nu hun krachten bundelen.

In de voorhoede

Inhoudelijk ligt het onderzoek nogal uiteen: waar het IISG sociale en economische geschiedenis onderzoekt op wereldschaal, concentreert het Huygens ING zich op Nederlandse geschiedenis, wetenschapsgeschiedenis en letterkunde en houdt het Meertens Instituut zich bezig met de Nederlandse taal en cultuur. Maar methodologisch voeren de instituten de afgelopen jaren dezelfde koers. Alle drie werken ze met data en ontwikkelen ze tools en websites. Misschien lopen de instituten van het HuC wel een beetje voorop, want taalwetenschappers en sociaal historici begonnen in de jaren tachtig al databases te bouwen.

KNAW-instituten hebben van oudsher een infrastructuur taak. “Die taak zetten wij voort”, zegt Henk Wals (IISG), “maar dan op een digitale manier.” Antal van den Bosch (Meertens Instituut) valt hem bij: “In de geesteswetenschappen is de stap naar het digitale helemaal geen moeilijke. De digitale manieren om informatie te coderen, liggen eigenlijk heel dicht bij hoe we al werkten. Misschien dat we daarom die digitaliseringsslag vrij snel gemaakt hebben. Zelfs sommige bèta’s zijn jaloers op ons. We zitten in de voorhoede.” Maar om interdisciplinair onderzoek mogelijk te maken, moet je er wel voor zorgen dat verschillende datasets op vergelijkbare manieren gekoppeld kunnen worden. Lex Heerma van Voss (Huygens ING): “Het Huygens ING



De drie directeuren, v.l.n.r.: Antal van den Bosch (Meertens Instituut), Lex Heerma van Voss (Huygens ING) en Henk Wals (IISG). foto Bart van Vliet

is gespecialiseerd in het maken van tekstedities van literaire en historische teksten. Honderd jaar lang wisten we precies hoe we dat moesten doen, maar door de digitalisering moet je nieuwe methoden en standaarden vinden. Daar zit een periode aan vast van experimenteren, waarin je niet weet of wat je doet wel het beste is. Nu kunnen we samen standaarden ontwikkelen waar meerdere disciplines mee kunnen werken.”

Ook een digitale heilige graal

Hoewel de digitalisering van data dus al langer gaande is, is het onderzoek met grote hoeveelheden data nog vrij nieuw binnen de geesteswetenschappen. Van den Bosch: “Vanaf het begin is er een harde kern geweest met voorlopers. Ook financiering vanuit KNAW en NWO heeft ervoor gezorgd dat alle instituten de digitalisering verder oppakten. Daar plukken we nu de vruchten van: er staat een generatie van jonge onderzoekers klaar die promoveert in digital humanities. Dat zijn onderzoekers die bijvoorbeeld ook kunnen programmeren. Dat wil overigens helemaal niet zeggen dat de traditionele manier van onderzoek ophoudt te bestaan. Het is geen of-of-

INTERVIEW

‘Eeuwenoude onderzoeksvragen beantwoorden met nieuwe digitale methoden’

verhaal, maar én-én. De twee kunnen elkaar versterken.” Zo heb je bij een geavanceerde liederendatabase nog steeds de blik van de liederenexpert nodig om de data te interpreteren en er een verhaal van te maken. En dat een goed geschreven verhaal ook voor historici nog steeds de heilige graal is, zoals Wals zegt, zal niemand verwonderen.

Meerwaarde van eHumanities

“De traditionele geesteswetenschapper had alleen een stukje papier en een pen nodig”,



KNAW Humanities Cluster

Het doel van het KNAW Humanities Cluster (HuC) is het versterken van vernieuwend, interdisciplinair, geesteswetenschappelijk onderzoek dat belangrijke wetenschappelijke en maatschappelijke thema's aansnijdt.
huc.knaw.nl

schetst Wals de ontwikkeling. “Nu krijg je een laboratoriumsituatie, waarin verschillende specialisten verschillende stukken van het onderzoek uitvoeren. We doen dat nu bijvoorbeeld op het thema migratie; aan migratiestromen liggen een heleboel harde data ten grondslag. En je kunt er vanuit allerlei perspectieven naar kijken: historisch, taalkundig, cultureel.”

Volgens Heerma van Voss is dit precies de meerwaarde van eHumanities: “Doordat je een veel grotere hoeveelheid gegevens kunt bevragen, krijg je betrouwbaardere kennis. Ook doordat je een dataset benadert vanuit verschillende disciplines. Het bevordert de interdisciplinariteit. Vroeger had je een heel duidelijke grens tussen taalkundigen, historici en letterkundigen en nu zie je dat ze allerlei zelfde technieken gebruiken.”

De digitale manier van onderzoek leidt tot meer onderzoek in teamverband, zegt ook Wals. Het HuC wil dit soort onderzoek stimuleren, maar tegelijkertijd moet het wel van onderaf ontstaan, benadrukt hij. “Vandaar ook dat we ervoor gekozen hebben om zo veel mogelijk op één locatie te gaan zitten, waar onderzoekers elkaar op de gang tegenkomen. We hopen dat zo mooie samenwerkingsverbanden ontstaan, op een natuurlijke manier.”

Internationale samenwerking

De digitalisering bevordert ook internationale samenwerkingsverbanden binnen vakgebieden, merkt Van den Bosch op. Een mooi voorbeeld is het onlangs toegekende digitaliseringsproject ISEBEL. In dit project worden de data uit de Nederlandse volksverhalenbank van het Meertens Instituut gekoppeld aan een Deense en Duitse verhalendatabase. “Dit soort netwerken zie je overal ontstaan.” Wals haalt tot slot een project aan van het IISG uit de jaren dertig – het precomputertijdperk – waarin men keek naar de ontwikkeling van lonen en prijzen. “Daarvan hebben we een prachtige gedrukte dataverzameling. In 2010 hebben we dit project weer opgepakt, omdat onderzoekers over de hele wereld vergelijkbare gegevens hadden verzameld, en omdat je nu digitaal die databases veel mooier en efficiënter kunt maken.” Het illustreert perfect hoe je met nieuwe digitale methoden voort kunt bouwen op eeuwenoude onderzoeksvragen.

huc.knaw.nl

In gesprek met Lotte Wilms, coördinator van het online KB Lab:

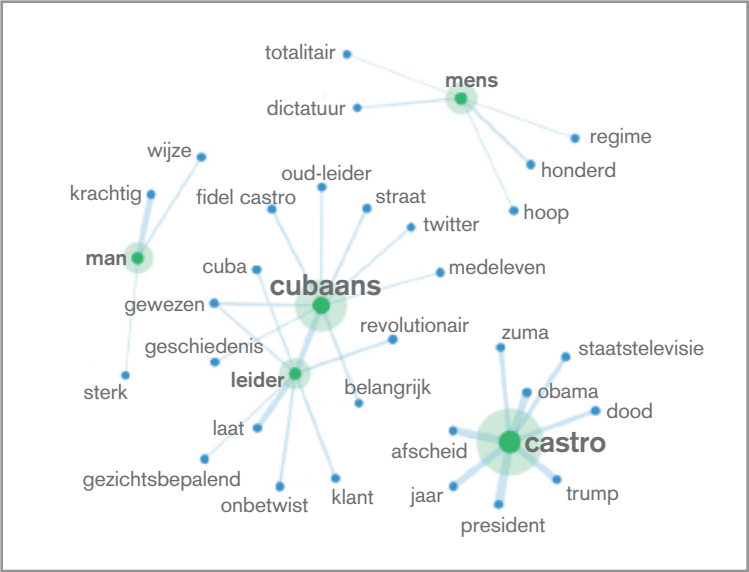
‘Tools en data uit digitale collectie beschikbaar’

Begin april lanceerde de Koninklijke Bibliotheek (KB) een volledig vernieuwd online lab. De website bevat vooral data en tools voor onderzoekers, maar de coördinator van het lab heeft nog meer plannen. *Steven Claeysens*

“Het KB Lab is in de eerste plaats bedoeld voor iedereen die wil werken met onze grote digitale collecties als datasets”, steekt Lotte Wilms van wal. “Dat kunnen computerwetenschappers zijn die een dataset nodig hebben om een bepaalde zoekinterface te bouwen, maar ook geesteswetenschappers die juist op zoek zijn naar een tool om te onderzoeken hoe een bepaald concept door de jaren heen in de gedigitaliseerde kranten verschuift. Daarnaast hopen we ook collega’s bij andere bibliotheken, archieven en musea te inspireren door te laten zien wat er allemaal kan met grote collecties gedigitaliseerd historisch materiaal.”

Voordelen nieuwe lab

Uiteindelijk hoopt Wilms dat het ook tot nieuwe toepassingen leidt: “We zoeken altijd nieuwe ideeën en projecten om samen met onderzoe-



Een van de beschikbare tools in het KB Lab is de Frame Generator. Door middel van topic modelling en co-occurrence vindt en visualiseert de Frame Generator patronen in een (grote) verzameling Nederlandstalige teksten, om zo inzicht te geven in de dataset. De applicatie is ontwikkeld in het kader van het KB-fellowship van prof. dr. Joris van Eij-natten in 2016. credits KB

kers en collega’s te kijken hoe we ons werkveld nog inspirerender kunnen maken.”

Wat biedt het nieuwe lab dat het oude niet bood? “Vooral meer context bij de tools en datasets, maar ook bij de activiteiten van het team achter het lab. Elke tool en elke dataset heeft een eigen pagina met

meer informatie over hoe iets is ontwikkeld, waar je de broncode of de data kunt vinden en waarvoor het tot nu toe is gebruikt. We hebben daarnaast ook een blog, een agenda waarin je kunt zien wat het lab organiseert, maar ook waar onze collega’s spreken over projecten of plannen van het lab. Je kunt je ook

inschrijven voor de nieuwsbrief om op de hoogte te zijn en blijven van al het lab-nieuws.”

Nog maar het begin

Voor Wilms is het allemaal nog maar het begin: “We willen nog meer samenwerken en we willen, hoe tegenstrijdig dat ook klinkt met een nieuwe website, meer offline te vinden zijn. We hopen echt een netwerk te kunnen bouwen van gelijkgestemden, mensen die het leuk vinden om samen te kijken naar andere mogelijkheden van bestaande applicaties of zin hebben om samen nieuwe technieken te ontwikkelen. Ook willen we zelf meer onderzoek gaan doen naar onze digitale collecties, zodat we beter begrijpen wat we nu eigenlijk aanbieden en wat ermee kan. Maar we kijken bijvoorbeeld ook naar de rol die bibliotheken spelen binnen de Digital Humanities.”

Iedereen kan helpen

Kunnen onderzoekers ook iets bijdragen aan het lab? “Altijd! Heb je een interessante vraag of een verfrissend plan, neem dan contact met ons op om samen te kijken of we iets op poten kunnen zetten.”

lab.kb.nl

Workshop erfgoeddata startsein voor digitale samenwerking

Wie in korte tijd een solide basis wil leggen om tot internationale cross-disciplinaire samenwerking te komen, doet er goed aan zich een week ‘te laten opsluiten’. *Heidi Berkhout*

Met welke methode kun je migrantenerfgoeddata verzamelen, duurzaam digitaal bewaren en bevraagbaar maken voor uiteenlopende groepen belangstellenden? Deze vraag stelden historica Marijke van Faassen en digitaal historicus Rik Hoekstra zichzelf. Het antwoord was even eenvoudig als doeltreffend: ze organiseerden de vijfdaagse workshop Migrant (R)e-collections.

Onconventioneel

Marijke van Faassen: “Het Lorentz Center stelde een locatie en organisatorische ondersteuning ter beschikking. We werden uitgedaagd om een onconventioneel programma te bedenken, bedoeld om bruggen te slaan tussen verschillende disciplines. Zo stonden in de workshops niet de sprekers centraal, maar het punt dat aan het eind van

de week bereikt moest zijn om verder te kunnen met het grotere doel: dat alle betrokkenen meerwaarde ontlenden aan het resultaat en dat veel mensen er gebruik van kunnen maken.” Rik Hoekstra vult aan: “Er was veel ruimte voor interactie in subgroepen en van elke deelnemer werd continu een actieve inbreng verwacht.”

Duurzaam toegankelijk

De workshop werd opgezet met Australische partners van de *Digital Humanities Group Western Sydney University* en *Curtin University* en het *Centre for Global Heritage and Development Leiden*. De workshop had dertig deelnemers uit Nederland, Australië, Engeland, België en de VS, die samen de vier domeinen geesteswetenschappen, computerwetenschappen, cultureel erfgoed en de migrantengemeenschappen vertegenwoordigden. Samen bogen ze zich over de vraag: welke methode brengt het verspreid bewaarde, schriftelijke en materiele, erfgoed van migranten zo bij elkaar, dat de onderzoeksgegevens digitaal duurzaam toegankelijk zijn? “De deelnemers kwamen tot de con-

clusie dat er behoefte is aan een ‘digitale marktplaats van data’, waar aanbod en vraag over cultureel erfgoed bij elkaar komen. We willen het erfgoed niet op een fysieke locatie bijeen brengen. Met zoveel uiteenlopende en internationaal verspreide collecties en even zoveel eigenaars, zou dat praktisch onhaalbaar zijn,” aldus Van Faassen. Hoekstra: “Ook werd duidelijk dat er geen overkoepelend systeem moet komen. De data blijven op zichzelf staan en gaan niet op in een nieuw systeem. Wel worden ze aan elkaar gekoppeld, waardoor de data doorzoekbaar worden.”

Start is gemaakt

Met gereedschappen als *Linked Data* en *Timbuctoo* zijn de basisvoorwaarden aanwezig, maar er is ook methodologische ontwikkeling nodig. Aan dit laatste wordt verder gewerkt met de workshoppartners aan de inhoudelijke kant en met het digital humanities ontwikkelteam van het Huygens ING, de Semantic Web-groep van de VU en Web Information Systems-groep van TU Delft aan de technische kant. *lorentzcenter.nl*



Waar kunnen migranten en hun nabestaanden terecht met hun ‘bagage’, met hun dagboeken, foto’s en verhalen? En aan welke voorwaarden moeten herinneringen voldoen om als databestand geschikt te zijn voor wetenschappelijk onderzoek? De workshop Migrant (R)e-collections was podium voor deze, en andere vragen. credits Jelk Kruk, SuperNova Studio, foto Image ‘Maxxi luggage’ CC BY-NC-SA 2.0 by Lounergie / Emiliano Dominic

KORT

Pilot-projecten testen CLARIAH-infrastructuur

Omdat de ingediende voorstellen van dusdanig hoge kwaliteit waren, heeft het bestuur van CLARIAH besloten niet twaalf, maar zestien research pilots te honoreren. De projecten, die dit voorjaar van start gingen, laten zien welke onderzoeksvragen er beantwoord kunnen worden met de data en tools die beschikbaar zijn in de digitale infrastructuur voor de geesteswetenschappen. Tegelijkertijd verkennen ze ook waar nog belemmeringen liggen in het gebruik van de infrastructuur, zodat deze kunnen worden aangepakt. De projecten zijn gelijkmatig verdeeld over de verschillende focusgebieden in CLARIAH die gebruikmaken van tekstuele, gestructureerde en audiovisuele data. (ER) *clariah.nl*

Maak kans op een data-subsidie

Tot 1 oktober kunnen onderzoekers bij DANS een aanvraag indienen voor een subsidie voor een klein dataproject. Deze zogenaamde KDP-subsidie wordt verstrekt aan projecten waarbinnen één of meer belangrijke datasets worden beschreven en toegankelijk worden gemaakt. De subsidie is vooral bedoeld voor deelgebieden waar nog niet tot nauwelijks aan data-archivering wordt gedaan, en waarbij de subsidie van DANS een stimulans kan zijn om dit juist wel te doen. De vorige ronde zijn negen projecten gesubsidieerd, variërend van de ontsluiting van paleoecologische datasets tot Oral Historyprojecten. (HB) *dans.knaw.nl*

Visiting Professor Richard Smiraglia

Richard Smiraglia, hoogleraar aan de *School of Information Studies* van de Universiteit van Wisconsin-Milwaukee, is specialist op het gebied van zogenaamde *Knowledge Organisation Systems* en auteur van meer dan driehonderd publicaties relevant voor de informatiewetenschappen. Dankzij een beurs uit het *KNAW Visiting Professors Programme* kan Smiraglia zich in 2017 en 2018 buigen over vragen rond de ontwikkelingen van een grootschalige onderzoeksinfrastructuur binnen de Nederlandse geestes- en sociale wetenschappen. Smiraglia’s kennis en ervaring zullen met name worden gebruikt om de diensten van DANS, zijn gastinstituut, verder te ontwikkelen. Uiteraard streeft DANS ernaar de expertise van Smiraglia te delen. Richard Smiraglia is in mei begonnen. (CH)

dans.knaw.nl

Onderzoekslab Media Suite biedt tools voor mediawetenschappers

Audiovisuele data via één lab

Begin april lanceerde CLARIAH de eerste versie van de Media Suite, een online onderzoeks-omgeving voor mediawetenschappers. De eerste gebruikers helpen ook met de verdere doorontwikkeling.

Erica Renckens

“Dit is een historisch moment”, vertelt een trotse Roeland Ordeman, een van de ontwikkelaars, tijdens de lancering van de Media Suite. “De eerste stappen hiervoor werden al in 2004 gezet in het CATCH-project en later voortgezet in CLARIN.”

Openbaar toegankelijk

Voortaan kunnen mediawetenschappers in een online omgeving grote hoeveelheden audiovisuele data verkennen, doorzoeken, annoteren, analyseren en visualiseren. De Media Suite bevat verschillende tools die al eerder beschikbaar waren, maar niet eerder waren deze onderling compatibel en openbaar toegankelijk voor onderzoekers.

Gebruikers van de Media Suite loggen via Surf in met hun institutionele account en krijgen dan toegang tot componenten waarmee ze zelf hun eigen tool kunnen samenstellen. In de praktijk blijken verschillende combinaties van componenten vaker voor te komen; de ontwikkelaars hebben deze alvast samengebracht in vier verschillende ‘recepten’. Hiermee kunnen onderzoekers eenvoudig collecties verkennen, doorzoeken en vergelijken.

Beter begrip

Na afloop van de lancering mailt Julia Noordergraaf, leider van het CLARIAH-werkpakket gericht op Mediastudies: “De eerste gebruikers zijn gelijk onze testers en co-ontwikkelaars. Alles in de Media Suite werkt, maar het is in feite nog een vrij kaal huis met alleen de basisvoorzieningen. Aan de verdere inrichting moet nog veel gebeuren. De eerste gebruikers waren tijdens de lancering zeker onder de indruk van de mogelijkheden, maar hadden ook nog enige moeite om hun weg te vinden in dit onderzoekslab ‘in aanbouw’. Een concreet aandachtspunt is de vormgeving: de kleuren in de tijdlijn zijn bijvoorbeeld voor kleurenblinden niet te onderscheiden. Een meer fundamenteel punt was de confrontatie met de ‘ruwe’ weergave van metadatavelden, die ver-



De eerste gebruikers van de Media Suite zijn de uitvoerders van de onlangs gestarte research pilots van CLARIAH. Zij mochten tijdens de lancering op 4 april jongstleden in het Nederlands Instituut voor Beeld en Geluid als eerste rondneuzen in de nieuwe onderzoeksomgeving. foto Erica Renckens

warrend bleek. Een grote collectie als die van Beeld en Geluid heeft bijvoorbeeld voor ‘datum’ al vijftien verschillende datavelden. Anderzijds zorgt deze confrontatie voor een beter begrip van de structuur van de datacollectie waarmee wordt gewerkt. De input van

gebruikers nemen we de komende anderhalf jaar mee in drie nieuwe versies van de onderzoeksomgeving, steeds met nieuwe functionaliteiten en verbeterd gebruikersgemak.”

mediasuite.clariah.nl

Generieke software voor geesteswetenschappers

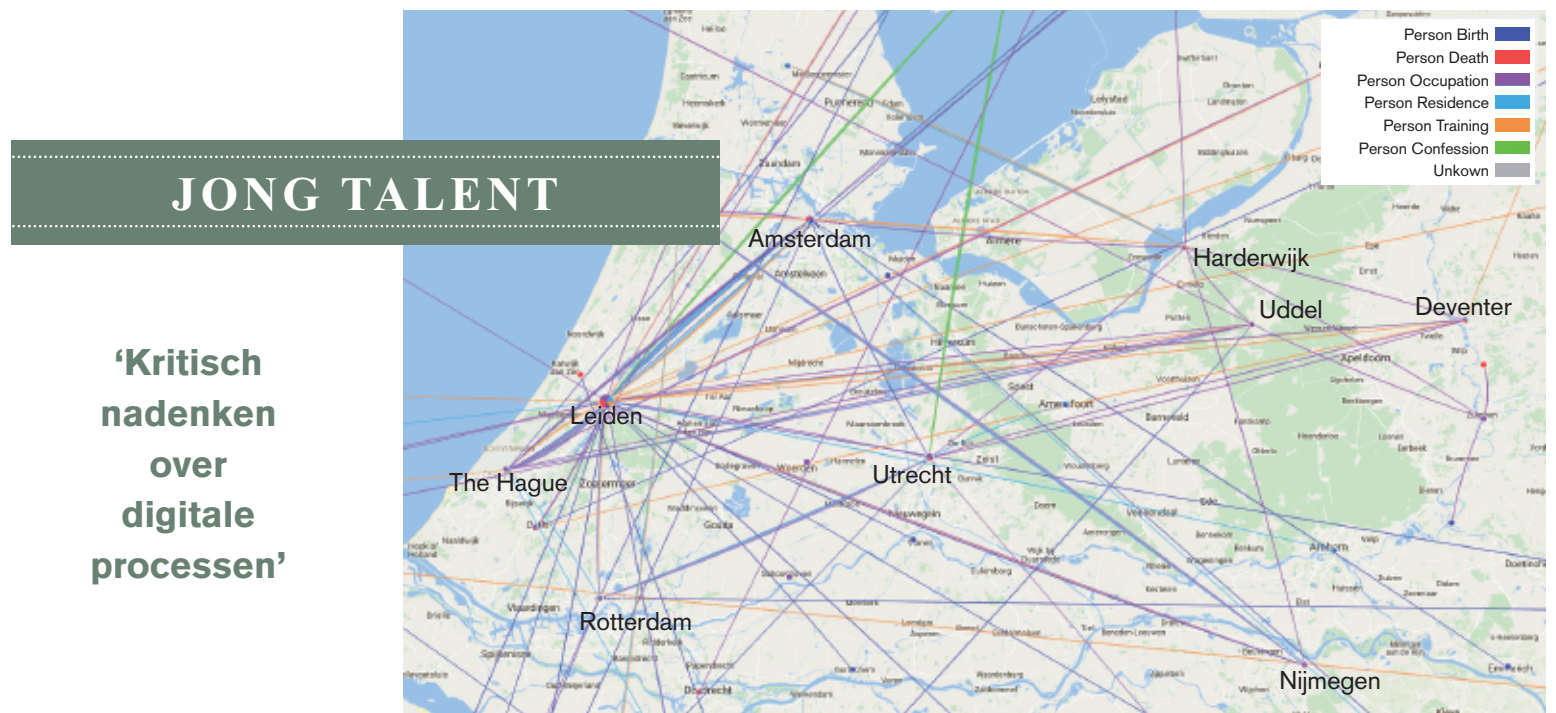
Draagvlak voor zelfgebouwde database

Op het kerkhof van Digital Humanities liggen al genoeg overbodige tools. Pim van Bree en Geert Kessels (LAB 1100) ontwikkelen generieke software voor geesteswetenschappers. Machteld Maris

“Laten we hier een misverstand uit de wereld helpen. Digital natives, de generatie geboren en getogen mediagebruikers, kunnen weliswaar hun status op Facebook aanpassen en weten hoe ze moeten Snapchatten, Whatsappen en Instagrammen. Maar ze zijn niet gewend om kritisch na te denken over digitale processen. Google is voor hen gewoon een vanzelfsprekende vraagbaak.” Kessels en Van Bree zeggen het al jaren: opleidingen, ook in de geesteswetenschappen, zouden in het eerste jaar moeten beginnen met lessen in digitale vaardigheden. Veel geesteswetenschappers zijn huiverig voor informatica. Zij schrijven liever een boek. Maar een onderzoeker anno 2017 moet toch op zijn minst een database kunnen bouwen.

Nodegoat

Hun bedrijf, LAB1100, is het enige softwareontwikkelingsbureau in Nederland dat zich specifiek richt op academisch onderzoek in de geesteswetenschappen. Kessels en Van Bree ontwikkelden in 2011 nodegoat, een web-based onderzoeks-



‘Kritisch nadenken over digitale processen’

omgeving voor wetenschappers. Node staat voor een knooppunt in een bepaald netwerk. Een boek, een brief, een persoon, een instituut. Goat, inderdaad, voor dat intelligente, nieuwsgierige en koppige dier.

Big Data is overal

Van Bree: “Tegenwoordig struikel je over de term Big Data. De focus ligt op enorme datasets. Beste voorbeeld daarvan is natuurlijk Delpher van de KB. Wij hanteren een bottom-up aanpak. De onderzoeker legt zelf, handmatig, een database aan in nodegoat. Onderzoekers kunnen

Geografische visualisatie gemaakt in nodegoat van biografische gegevens uit het project ‘Das Entschwinden der Erinnerung. Was heißt wissenschaftliches Vergessen-Werden im 18. Jahrhundert?’, uitgevoerd door dr. Tobias Winnerling, Heinrich-Heine-Universität Düsseldorf. Op nodegoat.net kunnen onderzoekers een eigen onderzoeksomgeving inrichten. credits dr. Tobias Winnerling, Heinrich-Heine-Universität Düsseldorf

hierin data verzamelen, ordenen, op verschillende manieren (zelfs interactief) visualiseren en openstellen voor anderen.”

Kessels: “Tijdens onze workshops leren deelnemers de basisbeginselen van het bouwen en gebruiken van een database. Als ze dat in de vingers hebben, kunnen ze vervolgens hun onderzoeksdata modelleren. Uiteraard is het van belang dat het aanleggen ervan in een vroeg stadium

gebeurt. Je moet er niet aan denken dat je al drie jaar onderzoek doet, en dan met terugwerkende kracht alle gegevens moet invoeren.”

Het 1+1=3 effect

Wie met een druk op de knop antwoorden zoekt, is bij LAB1100 aan het verkeerde adres. Ja, het aanleggen en bijhouden van de database is veel werk. Maar met de interface kun je sociale en geografische

verbanden zichtbaar maken die in tekstanalyse verborgen zouden blijven. Valorisatie is tegenwoordig onmisbaar in wetenschappelijk onderzoek. Wie de techniek van het datamodelleren in de vingers heeft, kan op verschillende manieren het proces en het belang van het onderzoek laten zien, en niet pas als het resultaat ervan in de boekwinkel ligt.

<https://nodegoat.net>

Zelf datasets aanmaken en delen

DataverseNL biedt netwerk van data-repositories

Online opslaan, delen en registreren van onderzoeksgegevens, gedurende de onderzoeksperiode tot aan de voorgeschreven termijn van tien jaar na de voltooiing ervan. Dat is DataverseNL. Marion Wittenberg

Sinds 2014 is DataverseNL een dienst van DANS. Elf universiteiten, twee hogescholen en een KNAW-onderzoeksinstituut maken er op dit moment gebruik van. Medio februari is de migratie naar 4.6 gereedgekomen, hierdoor is de userinterface gebruikersvriendelijker en ziet de website er mooier uit.

Maatwerk per instelling

Het principe achter DataverseNL is dat DANS de technische infrastructuur beheert, maar dat de participerende instellingen zelf verantwoordelijk zijn voor het gebruik en het datamanagement van hun DataverseNL-repositories. De invulling kan per instelling verschillen. Bij sommige instellingen kunnen onderzoekers zelf datasets aanmaken, maar worden deze pas gepubliceerd na controle door een datamanager. Bij andere instituten heeft de onderzoeker uitgebreidere rechten. Datasets worden aangemaakt in zogenaamde dataverses, zelfstandige eenheden waarin het databeheer



Het onderzoeksteam RiverCare. foto RiverCare

plaatsvindt. Per dataverse is het mogelijk om te variëren in gebruikersrollen, zoals administrator, dataverse curator, dataset creator of groeplid. Datasets kunnen gedeeld worden voordat ze zijn gepubliceerd, dit maakt DataverseNL met name zeer geschikt voor het delen van data tijdens onderzoek. De dienst is disciplineonafhankelijk.

De adviesraad bepaalt

De software waar DataverseNL gebruik van maakt, wordt ontwikkeld door het *Institute for Quantitative Social Science* (IQSS) van Harvard (VS). Voor Nederland ontwikkelt DANS, in nauwe samenwerking met IQSS, een aantal aanpassingen. Zo is

het voor Nederlandse gebruikers mogelijk om in te loggen met hun lokale account via SURFconext. Verder maakt DataverseNL gebruik van handles als Persistent Identifier. Momenteel wordt er gewerkt aan een automatische archiveringsinterface met het langtermijnarchief van DANS en 4TU.Datacentrum. Het beleid achter de ontwikkelingen van DataverseNL wordt bepaald door de adviesraad, waarin alle participerende instituten zitting hebben. Maakt uw instelling nog geen gebruik van DataverseNL, maar bent u geïnteresseerd in de mogelijkheden, neem contact met ons op. dataverse.nl

Rivieren vol data

In het NWO gefinancierde onderzoeksprogramma RiverCare werken promovendi, postdocs en junior onderzoekers van de universiteiten Twente, Delft, Wageningen, Utrecht en Nijmegen samen om de gevolgen van de aanpassingen aan het Nederlandse rivierenstelsel te onderzoeken. Er worden uiteenlopende gegevens verzameld waaronder data op het gebied van hydraulica, morfologie en ecologie. De RiverCare-onderzoeksgroep verkent momenteel hoe zij de DataverseNL-functionaliteit het beste kan inzetten voor de data behorende bij hun eerste publicaties. ncr-web.org/rivercare/about

GELEZEN

The Datafied Society Studying Culture through Data. Edited by Mirko Tobias Schaefer and Karin van Es
René van Horik



Omdat digitale data steeds belangrijker worden, zijn deze voor onderzoekers in toenemende mate een bron voor het bestuderen van de maatschappij. Deze *dataficatie* van de samenleving vergt nieuwe middelen om menselijk gedrag te begrijpen en te monitoren. 'The datafied society' beoogt een theoretische en methodologische toolkit te zijn voor onderzoekers die cultuur willen bestuderen op basis van digitale data. Het boek bespreekt de rol en bruikbaarheid van Big Data en introduceert methoden om deze te analyseren en te visualiseren. Ook politieke en ethische vragen over de verzameling, behandeling en presentatie van data komen aan bod. Het boek bevat 19 bijdragen onderverdeeld in vier secties: Studying Culture through Data, Data Practices in Digital Data Design, Research Ethics and Key Ideas in Big Data Research. De auteurs zijn afkomstig uit verschillende disciplines zoals media-studies, computerwetenschap en sociale wetenschappen.

nl.aup.nl/books/9789462981362-the-datafied-society.html

COLUMN

Kwetsbare data

De afgelopen jaren heb ik mij beziggehouden met christelijke jeugdliteratuur. Vooral met protestantse en katholieke jeugdverhalen over jodenbeking, want daar ging mijn proefschrift over. In totaal vond ik ruim tachtig van dergelijke jeugdverhalen. In sommige bekeringsverhalen gaat het er opvallend ruig aan toe. Kinderen die zich tot het christendom bekeren, worden door hun joodse ouders of grootouders bruut mishandeld en soms zelfs vermoord. In een populair katholiek verhaal wordt een jonge joodse jongen eerst uitgehongerd en vervolgens gekruisigd door zijn vader, met behulp van de rabbijn. Mede vanwege dit grove geweld was ik zeer geïnteresseerd in de receptie van deze jeugdverhalen. Daarom heb ik zo veel mogelijk recensies bij elkaar gezocht, wat een hele klus was. Een van de rijkste bronnen bleek uiteindelijk achterderug.nl, een particuliere website over christelijke jeugdliteratuur van Rens Strijbos. Op achterderug.nl staan de bibliografische gegevens van duizenden zondagsschoolboekjes, vaak samen met informatie over

de auteur, foto's en soms het complete jeugdboek in pdf-formaat. Daarnaast verzamelde Strijbos van duizenden christelijke jeugduitgaven de besprekingen, die eveneens op zijn site te vinden zijn. Van de door mij onderzochte jeugdverhalen over jodenbeking vond ik uiteindelijk ruim 140 besprekingen, in allerlei bronnen. Overigens wordt in geen van die recensies het excessieve geweld in deze verhalen genoemd, laat staan afgekeurd, maar dat terzijde.

Ik vind de site van Strijbos geweldig en stuur hem sinds enige tijd zo veel mogelijk aanvullingen, maar iedere site van een particulier is per definitie kwetsbaar. Bovendien is lang niet iedereen met achterderug.nl bekend. Voor bibliografische gegevens en de locatie van boeken kijken de meeste mensen in PiCarta of WordCat. De belangrijkste bron

voor informatie over jeugdboeken is het Centraal Bestand Kinderboeken (CBK). Deze databank bevat ruim 251.000 beschrijvingen van voornamelijk Nederlandstalige kinderboeken, tijdschriften en centsprenten, plus zo'n 65.000 beschrijvingen van boeken en tijdschriftartikelen over jeugdliteratuur.

Bij mijn weten bevat het CBK geen besprekingen. En in PiCarta staan volgens mij alleen maar besprekingen van de bibliotheekdienst. Dat zou, voor wat be-

treft duizenden christelijke jeugdboeken, in één klap kunnen veranderen door bij die boeken een link op te nemen naar achterderug.nl. Of nog beter: door de complete besprekingen op te nemen bij de titelbeschrijvingen. Ik heb aan Strijbos gevraagd hoe hij hierover denkt. Hij zou het toejuichen, want ook hij wil dat de informatie op zijn site zo makkelijk te vinden is, ook via andere webpagina's. Bovendien is hij zich bewust van de kwetsbaarheid van een site die door een particulier wordt onderhouden.

Zelf beschik ik ook over een aantal interessante aanvullingen. In samenwerking met Strijbos digitaliseerde ik de meeste jaargangen van onder meer de 'IDIL-gids voor jeugdlektuur' (1949-1970). Ook daarin zijn duizenden besprekingen te vinden. Die zijn eruit te vissen en

kunnen zonder al te veel moeite worden toegevoegd aan het Centraal Bestand Kinderboeken, PiCarta en/of WorldCat. Natuurlijk ben ik me ervan bewust dat bij dergelijke recente bronnen het auteursrecht een probleem kan zijn. Maar ik weet ook dat dit probleem in de praktijk zelden opspeelt en deels kan worden voorkomen met een disclaimer ('laat rechthebbenden zich melden'). Zeker is dat er relatief eenvoudig, zonder grote investeringen, een koppeling kan worden gemaakt tussen vele duizenden Nederlandse jeugdboeken en de besprekingen ervan. In mijn ogen zou je hiermee een goudmijn ontsluiten, want jeugdboeken zijn buitengewoon interessante bronnen voor maatschappelijke veranderingen. Uit eigen ervaring kan ik zeggen dat zo'n koppeling toekomstige onderzoekers in ieder geval heel veel tijd zou schelen.

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



foto Leo van Velzen