



E-DATA & RESEARCH

Jaargang 7 | 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: CBS, CentERdata, CLARIN-NL, DANS, Huygens ING en de Koninklijke Bibliotheek.

INHOUD

2 Interessante bijeenkomsten in de Agenda

3 Nederlands vrijgedrag in kaart gebracht

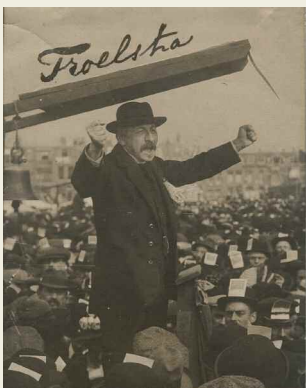
3 Krantencorpus online toegankelijk gemaakt

4 Deze datasets zijn sinds kort beschikbaar

5 De winnaars van de Dataprijs aan het woord

6 Wat is waar in onze digitale werkelijkheid?

7 Klopt het verhaal bij de beelden?



8 Geen geheimen meer voor wetenschappers

8 Ewoud Sanders is verslaafd aan e-data



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

Kwantitatief onderzoek naar orale cultuur

Zoekmachine herkent allerlei volksdeuntjes

Vroeger stond op menig schoorsteenmantel een bellenspeelklok. Die gaf niet alleen de tijd aan, maar speelde ook elk kwartier een deuntje. Een zoekmachine helpt deze melodieën herkennen.

Erica Renckens

Marieke Lefeber is onderzoekster aan het Meertens Instituut en conservator bij Museum Speelklok in Utrecht. Aan de hand van de bellenspeelklokken onderzoekt zij de muzikale smaak van de elite in de 18^e eeuw. Om de melodieën op de bellenspeelklokken te identificeren, maakte zij gebruik van de Nederlandse Liederenbank van het Meertens Instituut, sinds 2007 online. “De Nederlandse Liederenbank bevat metadata van ruim 150.000 liederen uit de middeleeuwen tot nu”, vertelt Martine de Bruin, projectleider Etnologie aan het Meertens Instituut.

Gelijkenissen en varianten

Ze werkte mee aan het CATCH-project Witchcraft waarin een melodieënzoekmachine voor de Liederenbank ontwikkeld is. “In Witchcraft hebben we gezocht naar een computationele methode om gelijkenis tussen melodieën vast te stellen. We wilden niet alleen exacte matches vinden, maar ook varianten van zo’n melodie.” In het project zijn eigenschappen die volksliedkundigen gebruiken om melodieën te herkennen, omgezet naar algoritmes waarmee de afstand tussen melodieën van volksliederen kan worden berekend.



Martine de Bruin (l) en Marieke Lefeber in Museum Speelklok te Utrecht foto Bart van Vliet

“Met behulp van Witchcraft heeft Marieke Lefeber diverse klinkende melodietjes kunnen identificeren”, aldus De Bruin.

In de Nederlandse Liederenbank zijn naast opnames ook muziekboekjes opgenomen, met daarin melodieën waarvan geen opnames

bestaan. Met behulp van crowdsourcing worden deze muziekboekjes getranscribeerd. De Bruin: “In onze nieuwe muzikeditor kan iedereen de muziek bij een gedeelte van de liedjes uit de Liederenbank bekijken, in transcriptie of in scan.”

VERVOLG OP PAGINA 2

Een veilig huis voor je data

Voor consumenten bestaan keurmerken om bijvoorbeeld gezonde producten te herkennen. Voor onderzoekers is het Data Seal of Approval steeds vaker het keurmerk voor veilig opgeslagen data.

Marika de Bruijne

Data kun je op verschillende manieren opslaan. Op je computer, op de universiteitserver, of in een digitaal nationaal archief. Sommige onderzoeken worden zelfs op YouTube gezet. De eerste optie, je eigen computer, is wel zo makkelijk. Maar veilig is het niet. Hoe weet je als onderzoeker waar je data wel goed kunt opslaan en hoe kun je dat beoordelen?

Het Data Seal of Approval (DSA), oorspronkelijk een initiatief van DANS, is een keurmerk dat verklaart dat de data-archieven die het keurmerk dragen, data veilig opslaan. Het keurmerk is momenteel aan acht data-archieven toegekend. In Nederland zijn dit EASY van DANS

en het Taal Archief van het Max Planck Instituut. De eisen van het keurmerk bepalen onder andere dat er speciale maatregelen moeten worden genomen om de veiligheid van de data, in de brede zin, te waarborgen. De dataopslag moet voor lange termijn geregeld zijn en procedures rondom gebruik van de data dienen transparant te zijn voor de buitenwereld.

Juist en integer

Ook is er veel aandacht voor het beschikbaar stellen van data aan andere academici. “Het DSA geeft aan dat onderzoekers erop kunnen vertrouwen dat de gegevens die ze ontvangen, grondig zijn beoordeeld op juistheid en integriteit”, vertelt Mary Vardigan, één van de board members van DSA en Assistent Directeur van ICPSR, het grootste digitaal archief in de sociale wetenschappen in de Verenigde Staten. “Er is aandacht geschonken aan de bruikbaarheid van de gegevens en het beschikbaar maken van meta-

data die nodig zijn voor het maken van analyses.”

Bekendheid groeit

Momenteel lopen zo’n dertien aanvraagprocedures en het is te verwachten dat het aantal uitgegeven keurmerken zich in 2013 minimaal zal verdubbelen. Dit betekent dat onderzoekers wereldwijd het DSA-logo steeds meer zullen tegenkomen. CentERdata is één van de organisaties die het DSA-keurmerk recentelijk heeft aangevraagd voor de data van het LISS panel. Deze data zijn zowel via een eigen website als via DANS toegankelijk voor alle academici. “Voor onze gebruikers is het belangrijk dat ze er van op aan kunnen dat onze data werkelijk betrouwbaar zijn, het keurmerk geeft dat op een simpele manier aan. Voor onszelf is het keurmerk een erkenning van alle moeite en tijd die we steken in de kwaliteit van onze data en documentatie”, vertelt Annette Scherpenzeel, projectmanager van het LISS panel.



Het DSA-keurmerk

Eind 2012 werd de eerste internationale conferentie over Data Seal of Approval gehouden. Er blijkt veel behoefte te zijn aan een universele en toegankelijke procedure rondom datacertificatie. Het doel voor de nabije toekomst is dan ook om de bekendheid van het DSA te vergroten onder zowel de archieven zelf als onder de dataproducenten en -gebruikers. “Ik hoop dat het keurmerk zich internationaal snel zal verspreiden, want het is een geweldige manier om op een transparante wijze te tonen dat je voldoet aan de normen van digitale duurzaamheid”, vindt Vardigan.

www.datasealofapproval.org

Dataverzameling via smartphones en tablets

Arnaud Wijnant

Op 6 november vond de NPSO-themamiddag over dataverzameling via smartphones en tablets plaats. Aan de hoge opkomst op deze middag was te merken dat het onderwerp momenteel erg populair is. De telefoons mochten aan blijven tijdens de presentaties. Niet om te bellen, maar om te twitteren over de middag. Hiervoor was speciaal de hashtag #npsomobile gereserveerd. Verschillende organisaties toonden hun innovatieve projecten met behulp van smartphones en andere mobiele apparaten. Veel van de getoonde mobiele applicaties maken gebruik van de talrijke sensoren die in een moderne smartphone te vinden zijn. Op deze manier kunnen onderzoekers het gedrag van mensen direct meten in plaats van indirect door er vragen over te stellen. Vooral de GPS-gegevens zijn populair voor het meten van verplaatsingen. Ook voor de respondent zijn er voordelen. Het verzamelen van data wordt vereenvoudigd door ge-



Dataverzameling door het LISS panel foto Hendri Adriaens

GEHOORD & BIJGEWOOND

bruik te maken van smartphones. Een voorbeeld is het Tijdsbestedingsonderzoek van het SCP, waarin respondenten in plaats van een dagboek bij te houden de dagelijkse bezigheden op een mobiel kunnen invullen. Smartphones en tablets worden al in verschillende onderzoeken gebruikt, maar de middag liet zien dat er nog veel meer mogelijk is.

www.npsa.net/node/151

ESTS: tekstedities voor mobiele apparaten

Peter Boot

In het uitgeven van historische en literaire teksten en bronnen is de afgelopen jaren veel veranderd. In zijn keynote op de conferentie van de European Society for Textual Scholarship (ESTS) onderkende Manfred Thaller vier generaties van digitale edities: de eerste publiceerde digitaal voorbereide edities nog op papier, de tweede generatie publiceerde elektronisch, vaak op CD-rom, maar de toen gekozen oplossingen zijn nu alweer verouderd. De huidige, derde generatie heeft voluit gekozen voor het internet als publicatieplatform, hoewel de hulpmiddelen om edities te maken qua gebruikersvriendelijkheid hopeloos achter de publicatiemogelijkheden aanhobbelen. En de vierde generatie komt eraan: moeten we tekstedities gaan maken voor mobiele apparaten? De ESTS-conferentie werd gehouden van 22 tot 24 november in Amsterdam. Huygens ING, zelf het resultaat van een recente fusie van een letterkundig en een historisch

onderzoeksinstituut gespecialiseerd in de uitgave van historische bronnen, had de conferentie georganiseerd om de historische en letterkundige tradities in bronuitgave dichterbij elkaar te brengen. In veel opzichten bleken de disciplines eigenlijk niet zoveel te verschillen: beide terreinen hebben te maken met dezelfde methodische en technische vragen. Hoe selecteren we ons materiaal? Hoe kunnen we 'leken' inschakelen bij het ontsluiten van bronmateriaal? Hoe richten we een technische infrastructuur in die rekening houdt met de veelheid aan interpretatiemogelijkheden van historische bronnen? Historici en letterkundigen kunnen op deze terreinen nog veel van elkaar leren.

ests2012.huygens.knaw.nl/

DE-conferentie in teken van 'Beter weten!'

Erica Renckens

Stichting DEN (Digitaal Erfgoed Nederland) organiseerde op 3 en 4 december samen met eenmalig partner CATCHPlus de tweejaarlijkse Digitaal Erfgoedconferentie (DE). Het thema van de conferentie was 'Beter weten', de vraag die centraal stond was: 'What can science do for you and what can you do for science?'. Hoewel sommigen vanwege de grote bezuinigingen in de erfgoedwereld bang waren voor een ietwat treurige sfeer tijdens de conferentie, bleek de stemming onder de aanwezigen juist positief en ambitieus. Nederland heeft een leidende positie in Europa en de rest van de wereld als het gaat om



foto Caskraker

samenwerking tussen erfgoed en wetenschap, met name op het gebied van digitalisering en infrastructuur. Dit blijkt bijvoorbeeld uit de rol van Nederland in projecten als CLARIN en Europeana. Dit betekent volgens de sprekers echter niet dat we op onze lauweren kunnen rusten. Zij pleitten bijvoorbeeld voor investeringen in het opleiden van een nieuwe generatie geesteswetenschappers die optimaal in staat zijn gebruik te maken van de digitale mogelijkheden. Daarnaast zouden erfgoedinstellingen zich meer onderling moeten verbinden, zodat data in een bredere context geplaatst kunnen worden. Hiervoor zijn open data onontbeerlijk.

www.den.nl

Workshop CLARIN over relatie met literatuurstudie

Peter Boot

Literaire teksten vragen om een andere benadering dan andere teksten. Of ze nu daadwerkelijk verschillen van andere teksten is de vraag, maar in elk geval worden er andere vragen aan gesteld. Een infrastructuur voor taaltechnologische hulpmiddelen, zoals CLARIN, moet daarmee

rekening houden. Die overweging was de reden voor Els Stronks (UU) en Karina van Dalen-Oskam (Huygens ING en UvA) om met steun van CLARIN een workshop te beleggen waar een aantal letterkundigen en een aantal ontwikkelaars van tools en infrastructuren bijeenkwamen om te praten over de mogelijkheden van (taal)technologie in de studie van literatuur. De workshop werd gehouden op 29 oktober in Den Haag. In de ochtend liet een aantal buitenlandse experts verschillende mogelijkheden zien. Tim Cole (VS) besprak de eisen die verschillende groepen wetenschappers stelden aan repositories van geesteswetenschappelijk bronmateriaal. Het vormen van een actieve community van wetenschappers is essentieel voor succesvolle repositories. Chris Meister (Hamburg) toonde zijn annotatietool CATMA, dat het interpretatieve werk van de geesteswetenschapper moet ondersteunen: meerlagige annotatie, samenwerking en elkaar tegensprekende bevindingen. Maciej Eder (Krakow) toonde hulpmiddelen voor stilistische analyse van teksten, op zoek naar kenmerken als bijvoorbeeld auteurschap of genre. Mike Kestemont (Antwerpen) en Karina van Dalen toonden een toepassing van zulke technieken speciaal op rijmwoorden. 's Middags werd gediscussieerd over de mogelijkheden die nieuwe technieken boden voor verschillende onderzoeksprojecten. Een duidelijke conclusie was dat de vorming van een community van onderzoekers die zich over deze vragen buigen een belangrijke stap voorwaarts zou zijn.

www.clarin.nl/event/446

VERVOLG VAN PAGINA 1

Volksdeuntjes

De gebruiker kan in de editor deze transcripties bijwerken en melodieën down- en uploaden.

Komend voorjaar

Inmiddels is in het CATCHPlus-project WitchcraftPlus het prototype van de melodieënzoekmachine doorontwikkeld tot een gebruiksvriendelijke toepassing. Komend voorjaar zal de zoekmachine voor iedereen online toegankelijk zijn. "Dan kan op melodieën uit de editor worden gezocht", vertelt De Bruin. "De gebruiker kan bij een melodietje klikken op een knop 'zoek gelijkende melodieën' of zelf een zoekopdracht invoeren door op een keyboard een melodie in te spelen of een MIDI-file te uploaden." Witchcraft staat aan de wieg van

een nieuwe richting onderzoek naar orale cultuur, waarbij gebruik wordt gemaakt van computationele methoden. In het E-lab Oral Culture werkt het Meertens Instituut samen met verschillende onderzoeksinstellingen, onder andere aan het project Tunes & Tales. Volksliederen werden generatie op generatie veelal mondeling doorgegeven, waardoor de verschillende varianten ontstaan die zoekmachine Witchcraft probeert te herkennen. In Tunes & Tales ontwikkelen onderzoekers een model dat deze variaties kan herkennen en voorspellen. Met deze kennis zal de zoekmachine in de toekomst beter varianten van een melodie kunnen herkennen.

www.liederenbank.nl
speelmuziek.liederenbank.nl

AGENDA

- 6 februari • Amsterdam**
Publish open access or perish?
De KNAW organiseert op 6 februari het symposium 'Publish open access or perish?'. Het doel van de middag is om de bewustwording rondom open access te verhogen, onder andere door controversiële onderwerpen met elkaar te bespreken.
www.knaw.nl/Pages/DEF/33/776.html
- 28 februari • Den Haag**
SURF Research and Innovation Event
Wetenschappelijk onderzoek in de 21^e eeuw is ondenkbaar zonder ICT-middelen en -methoden. Tijdens dit event tonen onderzoekers aan hoe multidisciplinair en data-intensief onderzoek wordt ondersteund en versterkt door en met ICT. Deze onderzoeksresultaten kunnen leiden tot nieuwe inzichten en grensverleggende onderzoeksvragen.
www.surf.nl/srie13

- 4 - 8 maart • Brussel (België)**
EU Science: global challenges & global collaboration
Beleidsmakers, wetenschappers en het bedrijfsleven komen samen tijdens deze door het Europese Parlement georganiseerde conferentie om verder te praten over Horizon 2020 en werkpakket 7.
www.globalsciencecollaboration.org/home
- 2 - 4 april • Washington (Amerika)**
Archiving 2013
Dit event is bedoeld voor conservatoren, archivariissen, bibliothecarissen en andere geïnteresseerden uit de culturele erfgoed gemeenschap om samen de uitdagingen op het gebied van digitale bewaring en beheer van dit erfgoed te bespreken.
www.imaging.org
- Vanaf 27 februari • Delft/Den Haag**
Cursus Data Intelligence
Een cursus bestemd voor medewerkers uit alle wetenschappelijke disciplines die meer willen weten over da-

- tamanagement en data-archivering. In 2013 wordt de eerste cursus georganiseerd op de woensdagen 27 februari, 27 maart, 24 april en 22 mei.
dataintelligence.3tu.nl
- 26 - 30 mei • Montpellier (Frankrijk)**
ESWC 2013
Tijdens dit event worden de laatste wetenschappelijke resultaten en technologische innovaties rond semantische technologieën besproken.
<http://2013.eswc-conferences.org/>
- 28 - 31 mei • Keulen (Duitsland)**
IAssist 2013
Het thema van de conferentie is Data Innovatie: de toegankelijkheid, zichtbaarheid en duurzaamheid. Dit thema weerspiegelt recente wereldwijde inspanningen van grote overheidsinstellingen tot kleine onafhankelijke onderzoekseenheden om gegevens meer open, toegankelijk en begrijpelijk voor iedereen te maken.
www.iassist2013.org

Vrijgebruiken in Nederland in kaart gebracht

Vrijers uit het verleden digitaal



Dankzij informanten is meer bekend over vrijers uit het verleden foto Chippix/Shutterstock

Afspraakjes maken, uitgaan en verkering krijgen. Voor velen, maar zeker voor historisch-demografen en gezinshistorici, interessante onderwerpen.

Douwe Zeldenrust en Hilde Bras

Al sinds 1931 stuurt het Meertens Instituut vragenlijsten naar informanten in Nederland en Vlaanderen. De vragenlijsten gaan over dialect, naamkunde en volkskunde.

Gemiddeld stuurden zo'n 1.000 mensen een ingevulde vragenlijst retour. Tegenwoordig worden de vragenlijsten digitaal verstuurd en ingevuld door het Meertens Panel bestaande uit ruim 5.000 informanten. De teller staat op 220 vragenlijsten. En deze worden de komende jaren gedigitaliseerd.

Op dit moment worden de antwoorden van vragenlijst nummer 40 uit 1971 gedigitali-

seerd. Lid van de Jonge Akademie van de KNAW en onderzoekster Hilde Bras: "Nummer 40 is een vragenlijst over vrijgebruiken. Het doel van deze vragenlijst was destijds om meer te weten te komen over verkering en verloving. De vragen gaan over de 'vrijer' en 'vrijster', de manier van kennismaken, de verkering en de verloving en bevat voor historisch-demografen en gezinshistorici zeer interessante gegevens over vrij- en verlovingsspatronen. Dergelijke gegevens kunnen op geen enkele andere wijze op systematische manier verkregen worden."

Ook data wordt gekoppeld

"We koppelen de culturele gegevens uit nummer 40 aan recent beschikbaar gekomen grote databestanden. Voorbeelden zijn individuele gegevens over levenslopen (Historische Steekproef Nederlandse Bevolking) en sociaal-economische en demografische gegevens op gemeenteniveau (Historische Databank Nederlandse Gemeenten). Zo worden eerdere bezwaren en tekortkomingen van de vragenlijsten deels ondervangen. Hiermee ontstaat er in wetenschappelijke zin meerwaarde boven wat er tot op heden op grond van de vragenlijsten is gedaan."

Douwe Zeldenrust, coördinator van de onderzoekscollecties van het Meertens Instituut, voegt hieraan toe: "In dit kader zijn er nog meer interessante vragenlijsten. Na de vragenlijst over vrijgebruiken wilden de onderzoekers destijds meer inzicht krijgen in huwelijksgebruiken. Vragenlijst nummer 43 uit 1973 gaat over die huwelijksgebruiken en kan ook voor hedendaags onderzoek significante gegevens bevatten. We verwachten de eerste resultaten halverwege 2013."

The power of the family

De digitalisering van vragenlijst 40 is een samenwerking tussen de Radboud Universiteit en het Meertens Instituut (KNAW). De data die dit oplevert, worden door Hilde Bras geanalyseerd in het kader van haar project 'The power of the family', dat door NWO gefinancierd is met een Vidi-subsidie vanuit de Vernieuwingsimpuls. Wilt u meer weten over data over daten? Neem contact op met Douwe Zeldenrust (douwe.zeldenrust@meertens.knaw.nl) of Hilde Bras (h.bras@let.ru.nl).

www.ru.nl/hlcs/the_power_of_the_family

Online krantencorpus maakt vergelijkend onderzoek mogelijk

De toon in kranten is de laatste decennia informeler geworden, zo is de algemene opinie. Maar om dat met cijfers te kunnen onderbouwen, is een groot geannoteerd corpus nodig. En dat is nu online toegankelijk.

Erica Renckens

"Uit eerder onderzoek bleek al dat Amerikaanse en Britse kranten informeler zijn geworden, met name in de tweede helft van de 20e eeuw", vertelt Kirsten Vis. Vis werkt als postdoc aan de Universiteit Utrecht en promoveerde in 2011 aan de VU op een soortgelijk onderzoek naar Nederlandse kranten. "Toen ik begon met mijn onderzoek was er geen materiaal uit 1950 digitaal beschikbaar, dus heb ik zelf pagina's gedigitaliseerd." Uiteindelijk leidde dit tot het VU-DNC-corpus (Vrije Universiteit - Dutch Newspaper Corpus), met daarin bijna twee mil-

joen woorden uit vijf grote Nederlandse landelijke kranten uit de perioden 1950/1951 en 2002. Dit corpus is sinds eind 2012 beschikbaar via de CLARIN-infrastructuur.

Toename subjectiviteit

Het corpus verschilt op twee punten essentieel van het krantenproject van de Koninklijke Bibliotheek (KB) waarin Nederlandse dagbladen online beschikbaar worden gemaakt. "Het project van de KB heeft van de grote landelijke kranten alleen op grote schaal artikelen uit De Telegraaf overgenomen", legt Vis uit. "Alleen van de periode rond de Tweede Wereldoorlog zijn ook artikelen uit andere kranten opgenomen." Het VU-DNC-corpus bevat artikelen uit De Telegraaf, Algemeen Dagblad, Trouw, de Volkskrant en NRC Handelsblad. Daarnaast is het VU-DNC-corpus voorzien van verschillende annotatielagen. Vis: "Naast woordsoort-

en lemma-informatie bevatten die lagen ook annotaties met subjectieve elementen die de mening, houding of emotie van de schrijver of geciteerde aangeven." Dat zijn bijvoorbeeld modale elementen ('mogelijk', 'zeker', 'hopelijk'), intensiverende elementen ('nogal', 'erg'), uitroepen ('wat mooi!') en voornaamwoorden ('ik', 'mijn', 'jouw'). "Bovendien heb ik de directe citaten als zodanig gemarkeerd. Zo kan onderscheid worden gemaakt tussen de tekst waarin de journalist zijn eigen woorden gebruikt en die waarin de woorden van een ander worden geciteerd."

Met behulp van het VU-DNC-corpus kon Vis alsnog de gestelde onderzoeksvraag beantwoorden. "Ik vond dat er inderdaad meer meningen, houdingen en emoties worden uitgedrukt in kranten uit 2002 dan in kranten uit 1950/1951."

Maar deze toename aan subjectiviteit was vooral te danken aan de

woorden van de geciteerden. "De journalist was helemaal niet subjectiever gaan schrijven, maar het gebruik van directe citaten was in de tijd verdubbeld. En in de citaten vond ik wel een toename van subjectiviteit."

Ijkpunt

Het VU-DNC-corpus biedt geesteswetenschappers niet alleen de mogelijkheid om vergelijkend kwantitatief onderzoek te doen naar landelijke dagbladen. Het corpus dient ook als ijkpunt voor OCR-postcorrectietools, waarmee spelfouten uit de automatische woordherkenning kunnen worden gecorrigeerd. "Na het digitaliseren zijn verschillende correcties uitgevoerd. De ongecorrigeerde en gecorrigeerde bestanden hebben we woord voor woord opgelijnd. Deze verzameling kan dienen voor toetsing en training van OCR-correctietools."

portal.clarin.inl.nl/vu-dnc/

KORT

KB lanceert de app 'Hier was het nieuws'

De KB lanceert de app 'Hier was het nieuws'. Met deze app kunt u overal in Nederland historische informatie oproepen over de plek waar u bent. De app van de KB is gebaseerd op de website Historische Kranten. Op deze website zijn miljoenen gedigitaliseerde krantenpagina's gratis toegankelijk. De app toont krantenartikelen tot 1940 en maakt gebruik van GPS. (HB)

www.hierwashetnieuws.nl



SINDS KORT BESCHIKBAAR

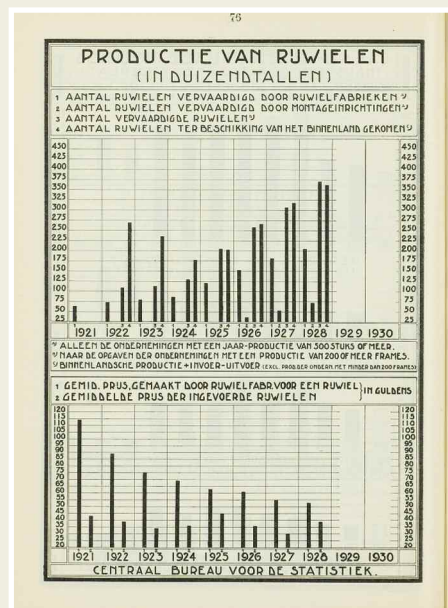
Dit overzicht toont databestanden die recent beschikbaar zijn gekomen bij het Centraal Bureau voor de Statistiek, CentERdata, Data Archiving and Networked Services en Huygens ING.

CBS

• Historische CBS-collectie online

Goed nieuws voor onderzoekers die de collectie historische-statistische publicaties van het CBS gebruiken voor hun onderzoek. In plaats van een bezoek aan het CBS zijn op termijn deze publicaties online te raadplegen. De collectie omvat zo'n 5 miljoen handgeschreven en gedrukte pagina's met tabellen en artikelen uit de negentiende en twintigste eeuw. Nadat eerder onder de noemer 'grote historische CBS-tellingen' de volks-, beroeps-, woning- en bedrijfs-tellingen online beschikbaar zijn gekomen, worden de publicaties in een meerjaren-project gedigitaliseerd en seriematig online beschikbaar gesteld.

www.cbs.nl/historisch



CBS gaat online: overzicht productie van rijwielen tussen 1921 en 1930 credits CBS

Ook sinds kort beschikbaar:

- Arbeidsongeschiktheidsuitkeringen, 2007 - 2012 (juni)
- Werkgevers Enquête Arbeid, 2008, 2010, 2012
- Huurenquête, 2012
- Buitenlandse zeggenschap over bedrijven in Nederland, 2010
- Nederlandse zeggenschap over buitenlandse bedrijven buiten de EU27-zone, 2010
- Survey Integratie Nieuwe Groepen, 2009
- Waarde onroerende zaken, 2012



Deze bestanden zijn beschikbaar via www.cbs.nl/microdata. Bezoek deze site of scan de QR code.

CentERdata

- Doelen stellen. Maas, J.; Ridder, D. de; Vet, E. de, Proactive Coping and Health Behaviour – februari/maart 2012

In februari en maart 2012 hebben Josje Maas, Denise de Ridder en Emely de Vet de studie 'Proactive Coping and Health Behaviour' afgenomen in het LISS panel. Het doel van de studie is om de validiteit en betrouwbaarheid van de Utrecht Proactive Coping Competence-schaal te bepalen. Dit is een schaal van 21 items die meet of individuen in staat zijn om realistische doelen te stellen, om mogelijkheden te ontdekken om te handelen naar doelen en om te gaan

met moeilijkheden. In deze studie zijn ook vragen gesteld over zelfwerkzaamheid, doeloriëntatie, zelfcontrole, toekomstig tijdsperspectief, preventieve coping, motivatie en intentie met betrekking tot gezond leven. De data uit het onderzoek zijn beschikbaar via LISS Data Archive.

Ook sinds kort beschikbaar:

LISS panel, nieuwe studies en data

- Maat, J. van de, The 'Don't Know' Option and the Outcomes of Opinion Polls, februari 2012
- Coffé, H.; Michels, A., Public Attitudes Towards Representative, Direct and Stealth Democracy, december 2011
- Tooren, M. van den; Rutte, C. G., Type D personality and job resources - Wave 1, januari 2012
- Horeni, O., Individual-level mental representations of complex decision problems - Wave 1, april 2010, en Wave 2, september 2010
- Gallego, A., Income Inequality, maart 2012
- Mulder, C. H.; Musterd, S.; Bertolini, L., Mobility in social networks - Wave 1, april 2009

LISS panel, nieuwe data

- Wilsem, J. van; Donker, A.; Leun, J. van der, Conventional and Computer Crime Victimization - Wave 2, februari 2010, en Wave 3, februari 2012
- Pieters, R.; Leenheer, J., Tilburg Consumer Outlook Monitor – september en december 2011, maart 2012
- CentERdata, Personality - Wave 5, mei/juni 2012
- CentERdata, Economic Situation: Housing - Wave 5, juni/juli 2012
- CentERdata, Family and Household - Wave 4, maart/april 2011; Wave 5, maart/april 2012

Immigrant panel, nieuwe studies en data

- Stupar, S.; Vijver, F. van de; Emotional Expressiveness - Part 1, januari 2012, en Part 2, februari 2012
- CentERdata, Family and Household - Wave 1, maart 2011
- Valchev, V.H.; Vijver, F. van de, Social-Relational Cluster, maart 2012
- Bender, M.; Chasiotis, A.; Vijver, F. van de, The function of autobiographical memory, oktober 2011

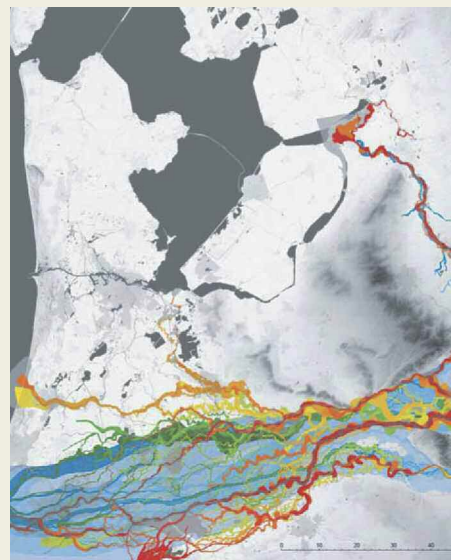


Deze bestanden zijn beschikbaar via www.lissdata.nl/dataarchive. Bezoek deze site of scan de QR code.

DANS

- Cohen, Dr. K.M. (Utrecht University); Stouthamer, Dr. E. (Utrecht University) (2012), Digitaal Basisbestand Paleogeografie van de Rijn-Maas Delta

Deze dataset en beknopte toelichting bevat de reconstructie van de ontwikkeling van het netwerk van stroomgordels in de Rijn-Maas delta in het Holoceen. Op de Universiteit Utrecht wordt sinds 1999 een GIS bijgehouden waarin de reconstructie van het riviernetwerk door de tijd wordt bijgehouden (K.M. Cohen, UU; Berendsen & Stouthamer, 2001; Berendsen, Cohen & Stouthamer, 2001; 2007). De hier gedeponeerde versie is geproduceerd in 2011-2012. Het betreft een volledige herziening en aanzienlijke uitbreiding van het eerder gekarteerde gebied. Een valorisatie subsidie van NWO-ALW, aangevuld met financiering



Rijn-Maas delta beeld UU

vanuit UU, TNO Geologische Dienst van Nederland en Deltares, hebben het herzieningsproject mogelijk gemaakt. De kaart volgt eerdere publieke versies (uit 2001) en interne versies van de UU (2003-2007) op. De dataset is kosteloos te downloaden in EASY.

urn:nbn:nl:ui:13-nqjn-zl

Ook sinds kort beschikbaar:

Archeologie

- Moesker, drs T.P., Hissel, M.E., Diachron UvA bv, 2012, Archeologisch onderzoek op bedrijventerrein Heeswijk-Dinther Retsel, gemeente Bernheze. AAC Projectenbureau - Diachron publicatie 54
- Blom, E., Feijst, L.M.B. van der, Veldman, H.A.P., ADC ArcheoProjecten, 2012-01, Beuningen Ewijk Plangebied keizershoeve I Opgraving en Begeleiding
- Diachron UvA bv; AAC Projectenbureau, september en oktober 2012, Wonen op de flank. Best, Dijkstraten

Archeologie gerelateerd

- Boon, Drs. P.J., Vries-Melein, Msc. M.M. de (2009), Mesopotamian cylinder seals: description and photographs of five seals from the NINO collection

Sociale Wetenschappen

- R. Eisinga, G. Kraaykamp, P. Scheepers, Radboud Universiteit: Religion in Dutch Society 2011-2012, DANS Data Guide
- J. M. Wicherts, Universiteit Tilburg: Cohort Differences in Big Five Personality Factors over a period of 25 years
- D.H. Zijssling, GION: Cohortonderzoek Onderwijsloopbanen van 5-18 jaar – COOL, Basisonderwijs, VO en VWO 2010/2011
- Centraal Bureau voor de Statistiek: Enquête Beroepsbevolking 2010 (beveiligd microbestand)

Historische Wetenschappen

- S. Leydesdorff en M. Huffener, Universiteit van Amsterdam en Stichting Sobibor: Interviews 'Late gevolgen van Sobibor'
- M. Marx, Universiteit van Amsterdam, PoliticalMashup 1814-2012

Taal- en literatuurwetenschappen

- N. van der Sijs, Stichting Vrijwilligers-netwerk Nederlandse Taal: Adriaen Koerbagh, Nieuw Woordenboek der Regten 1664 en Bloemhof 1668

Geodata

- Kadaster: TOP25 Raster, digitaal topografische bestand, september 2012



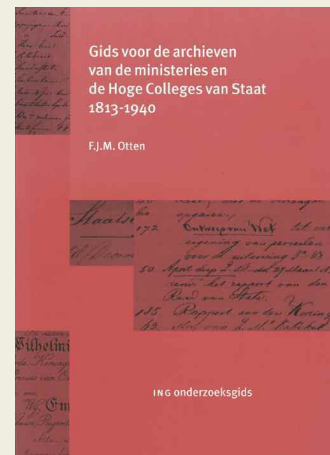
Via easy.dans.knaw.nl/ zijn deze bestanden beschikbaar. Bezoek deze site of scan de QR code.

Huygens ING

- Gids voor de archieven van de ministeries en de Hoge Colleges van Staat 1813-1940

In 2004 verscheen in de reeks ING-onderzoeksgidsen de 'Gids voor de archieven van de ministeries en de Hoge Colleges van Staat 1813-1940', geredigeerd door F.J.M. Otten, oud-rijksarchivaris bij de Rijksarchiefdienst. Deze is nu online raadpleegbaar op Historici.nl. Over elke instelling is dezelfde standaardinformatie opgenomen, namelijk over de geschiedenis van de organisatie, de ontwikkeling van het takenpakket, de wijze van de archiefvorming en de eigentijdse door de administratie zelf gevormde toegangen, de gedrukte bronnen en de belangrijkste institutionele literatuur. In de recente jaren zijn door het Nationaal Archief circa 6.600 archiefinventarissen op de website geplaatst. Vanuit de uitgebreide inhoudsopgave is het mogelijk op de openingspagina van de gedigitaliseerde toegang in het Nationaal Archief terecht te komen.

www.historici.nl/retroboeken/archiefgids_overheid



Gids voor de archieven van de ministeries en de Hoge Colleges van Staat 1813-1940 credits Huygens ING

Ook sinds kort beschikbaar:

- Gewestelijke Financiën ten tijde van de Republiek der Nederlanden (7 delen). De uitgave in boekvorm bevat een selectie en bewerking van de gegevens in de vorm van samenvoegingen van bedragen en herschikkingen van de posten in de bron. Het boek bevat tevens de inleiding en de toelichtingen op het materiaal. De database-uitgave staat reeds langer op de website
- Bronnen voor de geschiedenis der dagvaarten van de Staten van Zeeland 1318-1572 (3 delen). Naar analogie van het project 'Dagvaarten Holland' beoogt zijn Zeeuwse tegenhanger de reconstructie van de vergaderingen van de Staten van Zeeland en de daaraan voorafgaande bijeenkomsten van de landsheer met de Zeeuwse steden en van de Zeeuwse steden onderling. De editie bestrijkt de periode 1318-1572
- Oorkondenboek van het Sticht Utrecht tot 1301 (5 delen) Utrechtse oorkonden bewerkt door S. Muller Fz., A.C. Bouman, K. Heeringa en F. Ketner, uitgegeven in de periode 1920-1959



Deze publicaties zijn beschikbaar via www.historici.nl. Bezoek deze site of scan de QR code.

PoliticalMashup en OpenEarth winnaars Nederlandse Dataprijs 2013

Over terechte winnaars en knettergekke Kamerleden

'Prachtige projecten' en 'bijzonder inspirerende inzendingen'. Tijdens de uitreiking van de Nederlandse Dataprijs spraken de juryvoorzitters vol lof over alle ingezonden datasets.

Heidi Berkhout

Maar, er kan er maar één per categorie de winnaar zijn. E-data kijkt terug op de prijsuitreiking met Maarten Marx (PoliticalMashup) en Mark van Koningsveld (OpenEarth), winnaars van respectievelijk de Dataprijs humaniora en sociale wetenschappen en de Dataprijs exacte en technische wetenschappen 2013.

De lovende woorden van juryvoorzitter prof. dr. W.W. Mijnhardt staan in het geheugen gegrift van Maarten Marx. "Hij noemde PoliticalMashup 'omvangrijk, een grote variëteit aan bronnen, publiek toegankelijk en immens maatschappelijk relevant'. Echt mooie complimenten. Ik had niet verwacht de prijs te mogen winnen, ik was oprecht verrast", aldus een trotse Maarten.

Ook over OpenEarth werd enthousiast gesproken. Juryvoorzitter prof. dr. ir. J.Th. Fokkema verwoorde het als volgt: "Data verzamelen is nog maar het halve werk. Het is zeker zo belangrijk de data te rubriceren en te comprimeren. Op deze manier kunnen de data gedeeld worden met vakgenoten en andere geïnteresseerden. Het verzamelen en beschikbaar stellen van onderzoeksdata door OpenEarth is van uitstekend niveau en is daarom de terechte winnaar van de Dataprijs."

Dr. ir. Mark van Koningsveld (Van Oord/TU Delft) is samen met dr. ir. Gerben de Boer (Deltares/TU Delft) initiatiefnemer van OpenEarth. Mark glundert nog steeds bij de vraag waarom hij zelf denkt dat OpenEarth heeft gewonnen. "OpenEarth wordt veel gebruikt. Partijen uit de wetenschap en het bedrijfsleven leveren bovendien actief aan. Dit heeft, denken wij, ook een rol gespeeld bij het winnen van de prijs. Vrijwel dagelijks deelt onze community data, modellen en tools met elkaar. Een uniek aspect van OpenEarth, iets wat je op deze schaal zelden bij andere data- en kennismanagementinitiatieven ziet."

Oorsprong van het idee

"PoliticalMashup is ontstaan uit onvrede over de zoekinterfaces voor de Handelingen van de Tweede Kamer." Maarten vertelt verder aan de hand van een voorbeeld: "Wilders noemde eens een minister 'knettergek'. Was de ophef die ontstond over deze uitspraak terecht? Gebruikte Wilders als eerste deze term? Of was iemand hem voor, en wanneer was dat? Deze vraag bleek vrijwel onmogelijk te beantwoorden. En dat terwijl al sinds 1848 van elk in de Kamer gesproken woord wordt aangegeven: wie het zegt, in welke hoedanig-



*Vlnr: Maarten Marx, Thijs Damsma, Gerben de Boer en Mark van Koningsveld
foto Bart van Vliet*

heid, en wanneer." Maarten vervolgt: "Het probleem is dat de informatie er wel is, maar slechts beschikbaar in pdf-formaat. Alleen een (redelijk ingevoerd) mens kan dit lezen. In onze PoliticalMashup-dataset kan een computer de gegevens ook begrijpen. En kunnen gebruikers dus geavanceerde zoek- en analysevragen stellen. Inmiddels zijn alle Handelingen opgeknipt in zinnen en die zinnen zijn ontleed. En wat blijkt? Wilders was niet de eerste die de term 'knettergek' gebruikte. Hij was wel de eerste die die eigenschap aan een minister toekende. Eerdere sprekers werden zelf knettergek, of stelden dat het volk ergens knettergek van werd."

Interessant voor gebruikers

Ook OpenEarth is interessant voor gebruikers. Mark: "Voor, tijdens en na het uitvoeren van waterbouwkundige projecten wordt veel data verzameld. Denk aan gegevens over bijvoorbeeld stroming, golven, weer en wind, waterkwaliteitsmetingen. Dit soort gegevens zijn van groot belang voor bijvoorbeeld het maken van projectrapportages, het optimaliseren van productieprocessen en het doen van onderzoek. Daarnaast worden veel data gegenereerd

in de vorm van model-output. Het verzamelen, analyseren en visualiseren van deze gegevens kost veel tijd en vaak gebeurt er niets meer mee zodra een project is afgerond. Dat kan anders, dachten wij. Door informatie te delen, kan ieder project voortbouwen op de resultaten van voorgaande projecten. Gebruikers krijgen met OpenEarth in dezelfde tijd meer voor elkaar."

Positieve reacties

De winnaars hebben veel positieve reacties op de prijs ontvangen. "Uiteraard felicitaties, zelfs van de voorzitter van het college van bestuur van de UvA", aldus Maarten. Hij vervolgt: "Als onderzoeker steek je enorm veel tijd in een goed bruikbare dataset. Dat levert eigenlijk geen 'meetbare' opbrengst op in termen van artikelen of citaties. Het winnen van de Dataprijs, en zelfs al de nominatie daarvoor, onderstreept het belang van je werk. Dat voelt goed."

Ook voor Mark en zijn collega's geeft de prijs erkenning. "Over de jaren heen hebben we veel vrije tijd gestoken in de ontwikkeling en uitwerking van onze ideeën. Het winnen van de Dataprijs geeft aan dat we de goede weg zijn ingeslagen. Het geeft de voldoening dat alle geïnvesteerde tijd iets tastbaars, de Dataprijs, heeft opgeleverd. Bovendien merk je binnen onze organisaties dat externe erkenning van organisaties als DANS en 3TU.Datacentrum een belangrijke steun in de rug betekent."

Ideeën verder verspreiden

Mark: "OpenEarth is een standaard workflow. We zijn gebaat bij het vastleggen en verspreiden van protocollen en voorbeelden. We hopen dat het prijzengeld ons verder helpt alle

Dataprijs humaniora en sociale wetenschappen PoliticalMashup

maakt grote hoeveelheden politieke data digitaal doorzoekbaar.

• Maarten Marx

Dataprijs exacte en technische wetenschappen OpenEarth

is het alternatief voor het omgaan met data, modellen en tools in waterbouwkundige vraagstukken.

• Thijs Damsma, Gerben de Boer en Mark van Koningsveld

De Dataprijs bestaat uit €7.500, een sculptuur (zie foto) en een bon voor een geheel verzorgd diner.

verzamelde ideeën verder te standaardiseren, te documenteren en breder te verspreiden. In plaats van vanaf de zijlijn van ons dagelijks projectwerk kunnen we het nu op de manier oppakken zoals wij die voor ogen hebben." Ook Maarten heeft al een idee over de besteding van het geld. "We investeren in de ontwikkeling van de 'politieke ngram viewer', een geweldige tool om in de parlementsdata te grasduinen." Mark sluit af: "We nodigen onderzoekers in de waterbouw van harte uit om onze repositories te bezoeken. Voor veel van de vraagstukken waar je als (beginnende) onderzoeker tegenaan loopt, zijn al oplossingen beschikbaar. Met OpenEarth maak je als het ware een vliegende start. Bovendien biedt het werken binnen de OpenEarth community diverse voordelen. Er is een grote bereidheid om elkaar te helpen, om elkaars codes te verbeteren. Hierdoor kun je uiteindelijk beter werk leveren en vaak nog op een efficiëntere manier ook."

search.politicalmashup.nl
urn:nbn:nl:ui:13-3n38i9
publicwiki.deltares.nl

In de voetsporen van...

De Nederlandse Dataprijs, uitgereikt door 3TU.Datacentrum en DANS, geeft waardering aan onderzoekers die data bij elkaar brengen, documenteren en toegankelijk maken. In 2014 worden de Dataprijzen opnieuw uitgereikt. Meer informatie op www.dans.knaw.nl.

Iedereen een Digital Author Identifier (DAI)

Heidi Berkhout

Onderzoekers kunnen hun online vindbaarheid vergroten door gebruik te maken van de DAI. DAI staat voor Digital Author Identifier. De DAI is een uniek nummer voor elke auteur met een aanstelling of een andere relevante verbintenis aan een Nederlandse universiteit of onderzoeksinstituut.

Door het gebruik van dit nummer, bijvoorbeeld door deze bij het deponeren van data in EASY toe te voegen, wordt bij het hergebruiken en citeren van data altijd gerefereerd naar de juiste persoon. Verwarring door diverse schrijfwijzen van namen en identieke auteursnamen is verleden tijd.

Benieuwd naar uw DAI? Via de onderzoeksportal NARCIS kunt u uw DAI vinden op de persoonspagina's. www.narcis.nl

Eén ingang voor erfgoedcollecties

Esther Boeles

Onlangs heeft de Stichting Academisch Erfgoed (SAE) de website academischecollecties.nl gelanceerd. De nieuwe website biedt voor het eerst de mogelijkheid in één keer erfgoedcollecties van alle negen SAE-universiteiten te doorzoeken.

De objecten variëren van negentiende-eeuwse instrumenten tot anatomische preparaten, en van luchtfoto's tot mummieportretten uit het begin van de jaartelling. Een rijke bron voor geïnteresseerden in het academisch erfgoed en een ontdekkingsreis voor iedereen die nog niet wist dat deze waardevolle collecties bestonden. De betrokken universiteiten voegen elk half jaar content toe.

www.academischecollecties.nl



Lustrum AVSV 1927, studentes tijdens de feestelijke optocht
foto UvA Bijzondere Collecties

DANS gebruikerspanel

Wilt u meedenken over de diensten van DANS? Ga dan naar www.dans.knaw.nl en word lid van het gebruikerspanel.



Julia Noordegraaf foto Bart van Vliet

Wat is waar in onze digitale werkelijkheid?

Julia Noordegraaf, hoogleraar Erfgoed en Digitale Cultuur aan de Universiteit van Amsterdam, heeft goede ideeën. Een indruk van een inspirerend gesprek met een bevlogen vrouw. Heidi Berkhout

“Een archeoloog publiceert tegenwoordig niet meer alleen op papier. Kennis wordt verwerkt in driedimensionale constructies van vondsten. Dat vind ik fascinerend. Het zorgt ervoor dat je een veel beter inzicht krijgt in de oorspronkelijke samenhang van individuele vondsten. Deze moderne wijze van wetenschap bedrijven, is ook een vorm van publiceren. Maar deze stelt nieuwe vragen over wetenschappelijke betrouwbaarheid: Hoe doe je bij een 3D-reconstructie aan bronnenkritiek? Hoe annoteer je? Waar laat je je voetnoten?”

Erfgoed anders beleven

Aan het woord is Julia Noordegraaf. In haar kamer op de tweede verdieping van een van de gebouwen van de universiteit spreekt ze vol passie over haar ideeën. En dat zijn er voldoende. Met als gezamenlijk vertrekpunt het samenbrengen en bevorderen van onderzoek naar het hergebruik en de betekenis van digitaal erfgoed. Ook de invloed van digitalisering op de perceptie en waardering van cultureel erfgoed speelt voor haar een belangrijke rol. “Digitalisering binnen onze sector zorgt ervoor dat erfgoed bereikbaar is geworden voor velen. Kijk naar het Art Project van Google. Hiermee kun je kunstwerken van dichtbij en in meer detail bekijken dan in het echt met het blote oog mogelijk is. Dit zorgt ervoor dat je erfgoed op een andere manier kunt

beleven. Maar het zorgt ook voor het ontstaan van verschillende versies van ons erfgoed: populaire reconstructies bestaan naast wetenschappelijke versies. We moeten samen goed nadenken over het onderscheid tussen deze twee. Anders vervaagt ons idee over wat de werkelijkheid het meeste recht doet. De vraag ‘Wat is waar in het online domein?’, is een vraag die mij dan ook enorm boeit.

Hulp van de crowd

De crowd speelt voor de wetenschap een steeds grotere rol. Bij crowdsourcen helpen vrijwilligers bijvoorbeeld trefwoorden aan erfgoedobjecten toe te kennen. De vrijwilliger, de leek, kiest woorden op basis van wat hij ziet en hoort. Een expert kiest trefwoorden op basis van professionele standaarden. Interessant is het, om te onderzoeken of de combinatie leidt tot andere vormen van kennis. YouTube is een goed voorbeeld van crowduration. Iedereen kan zijn of

haar video's onderbrengen in een digitaal archief. Bij crowdconservation helpt de community de digitale data toegankelijk te houden. Bijvoorbeeld door het schoonhouden van databases. Dit zijn drie voorbeelden van hoe de crowd steeds meer betrokken is bij het beheren en ontsluiten van ons erfgoed. Het doet iets met onze identiteit, het maakt je onderdeel van de virtuele wereld. Daar kun je van alles van vinden, maar het heeft hoe dan ook gevolgen. En welke dit zijn, dat vind ik interessant om te onderzoeken. De antwoorden zorgen ervoor dat we

Erfgoedlab en CCHI

Momenteel richt Julia Noordegraaf binnen de UvA het Erfgoedlab op. Dit wordt een plek waar onderzoek in opdracht kan worden verricht. Daarnaast organiseert het 'Centre for Cultural Heritage and Identity' seminars, waar vraagstukken zoals genoemd in dit artikel worden besproken. Kijk voor meer informatie op www.cchi.uva.nl.

betere keuzes kunnen maken in hoe we ons erfgoed in de toekomst toegankelijk willen houden.

En als we niets doen?

Het Art Project kent een enorme toegevoegde waarde, maar vervangt niet de beleving die je ervaart als je voor een echt schilderij staat. Iemand live horen vertellen, is altijd anders dan een opname beluisteren. Maar erfgoedinstellingen die online niet bestaan, worden niet meer gevonden. De vraag is hoe erg dat is: de relevantie van historisch materiaal is altijd lastig te beargumenteren,

tot je wilt weten hoe iets precies zat. Dan wil je toegang tot dat cultureel erfgoed en de garantie dat wat je dan vindt, een betrouwbare weergave van de historische werkelijkheid is. Erfgoed is een terrein met veel emotie. Deze emotie bepaalt ook onze waardering. En dus is het des te belangrijker om de feitelijke context ook goed voor ogen te houden. Dit is het belang van standaarden, van samenwerken, van afspraken. Als we dat belang na blijven streven, kunnen we blijven toetsen wat waar is.”

www.cchi.uva.nl

Over crowdsourcen, crowduration en crowd-conservation

KORT

DBNL en KB in 2013 onder één dak

De twee grootste gedigitaliseerde archieven van kranten, boeken en tijdschriften in Nederland komen onder één dak. In 2013 wordt de DBNL namelijk ondergebracht bij de Koninklijke Bibliotheek. DBNL staat voor de Digitale Bibliotheek voor de Nederlandse Letteren. DBNL, een Nederlands-Vlaamse cultuurhistorische Bibliotheek, bestaat uit volledige, gedigitaliseerde literaire teksten, secundaire literatuur en aanvullende informatie zoals biografieën, portretten en verwijzingen naar andere bronnen. Het besluit is door het Bestuur van de Stichting DBNL in samenspraak met de Nederlandse Taalunie en de KB genomen. De betrokken partijen zijn ervan overtuigd dat deze ontwikkeling een positief effect zal hebben op de dienstverlening via de DBNL-website aan het wetenschappelijke en het algemene publiek. (HB) www.dbnl.org

Archieven Beeld en Geluid opengesteld

Het Nederlands Instituut voor Beeld en Geluid heeft zich als datacentre aangesloten bij CLARIN-nl. Voortaan zijn hiermee de data uit de archieven van het instituut beschikbaar voor geesteswetenschappers die gebruik maken van de CLARIN-infrastructuur. Naast Beeld en Geluid stelden eerder al de Koninklijke Bibliotheek, het Meertens Instituut, DANS, INL, Max Planck Instituut en het Huygens Instituut hun digitale archieven open voor de wetenschap. De meeste van deze datacentra zijn ook bereid data en tools van derden in beheer te nemen, zodat deze toegankelijk worden voor andere onderzoekers. (ER)

Oorlogsbronnen.nl verder uitgebreid

Onlangs zijn een aantal collecties aan de zoekmachine oorlogsbronnen.nl toegevoegd. Oorlog in Blik is één van die collecties. Oorlog in Blik presenteert materiaal van filmers uit de periode 1933-1950. Ook de oorlogsmonumenten in Nederland, interviews met veteranen over hun in ervaringen in de Tweede Wereldoorlog en de collectie van muziekwebsite musicforce.org zijn nu via de zoekdienst te vinden. Oorlogsbronnen.nl is een zoekmachine specifiek gericht op de Tweede Wereldoorlog in Nederland en in de voormalige koloniën. De zoekdienst biedt onderzoekers en andere geïnteresseerden de mogelijkheid om met één muisklik door twee miljoen objecten (foto's, tekeningen, video's, dagboeken, archiefstukken, audiofragmenten, interviews) uit tientallen archief-, bibliotheek- en museumcollecties te grasduinen. (HB) www.oorlogsbronnen.nl

Fotodatabank van onderzoeker Martijn Kleppe toont curieus gebruik beeldmateriaal aan

Klopt het verhaal bij de foto?

Onze hersenen verwerken voortdurend de beelden die we zien en die ons besef van de wereld beïnvloeden. Maar hoe wordt bepaald welke beelden we zien? En klopt het verhaal wel bij de beelden?

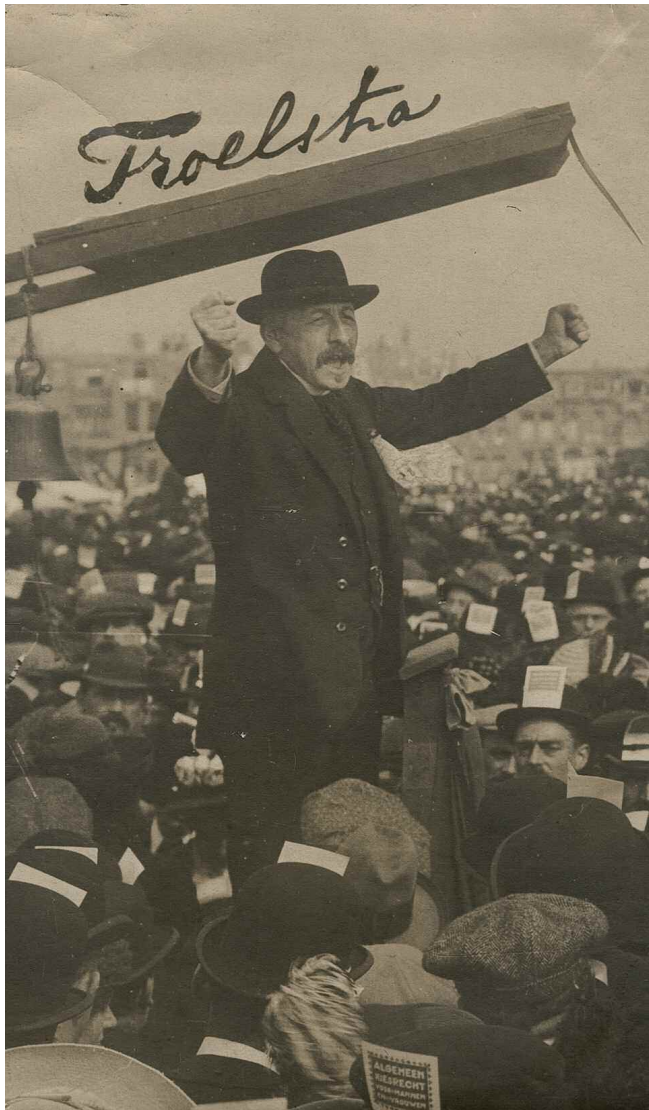
Marika de Bruijne

Iemand die zich met deze vragen bezig houdt, is Martijn Kleppe van de Erasmus Universiteit Rotterdam. Voor zijn promotieonderzoek heeft Kleppe foto's in Nederlandse geschiedenisboeken geanalyseerd. Door ze in een fotodatabank op te nemen, heeft hij kunnen onderzoeken welke foto's, door de jaren heen, het vaakst gebruikt worden. Zijn conclusie is dat dezelfde foto's soms aan verschillende verhalen worden gekoppeld. Een voorbeeld hiervan is een foto van Pieter Jelles Troelstra (zie foto en bijschrift). Een dergelijke manier van omgaan met foto's is op zijn minst curieus, omdat lezers foto's kunnen gebruiken als bevestiging van een verhaal.

Fascinatie voor beelden

“Tijdens mijn studie liep ik stage bij World Press Photo. Daar raakte ik gefascineerd door de manier waarop beelden onze wereld beïnvloeden”, vertelt Kleppe. “Elk jaar kiest World Press Photo uit honderdduizend foto's dé foto van het jaar. Deze foto verschijnt vervolgens in alle media, miljoenen mensen zien deze foto. De vraag is: waarom is die foto relevanter dan alle andere om de wereld van nu te beschrijven?”

Dat het proces van foto's beschrijven en typeren niet eenvoudig is, heeft Kleppe ook tijdens zijn onderzoek geconstateerd. Om de fotodatabank op te zetten, moesten ruim 5.000 foto's verrijkt worden met metadata die beschrijven wat er op de foto staat, zoals: het onderwerp, de personen, de gebouwen, het



1912 of 1918?

Dezelfde foto wordt soms aan verschillende verhalen gekoppeld. Zoals deze foto van Pieter Jelles Troelstra, waarop hij een gepassioneerde toespraak houdt tijdens een demonstratie voor de invoering van algemeen kiesrecht op 17 september 1912 in Den Haag. Deze foto wordt in schoolboeken echter vaak gebruikt om zijn mislukte revolutie van 1918 te verbeelden.

credits Corn. Leenheer, IISG

jaartal en bijvoorbeeld de locatie. Drie mensen hebben hier bijna een jaar aan gewerkt. “Beelden moeten nog steeds voornamelijk handmatig gecatalogiseerd worden. Hier zit enorm veel menskracht in. Bij het Nederlands Instituut voor Beeld en Geluid bijvoorbeeld worden alle uitzendingen van de publieke omroep nog voornamelijk handmatig omschreven om het beeldmateriaal vindbaar te maken”, zegt Kleppe. “Je kunt je nauwelijks voorstellen hoeveel werk dit is.”

Nieuwe technologieën, zoals automatische beeldherkenning, worden momenteel onderzocht. Eén project op dat gebied is Access to

Audiovisual Archives (AXES). Kleppe werkt hieraan mee door te kijken naar de manier waarop mensen zoeken naar beeldmateriaal en hoe nieuwe technieken hen daarbij kunnen helpen. Een oplossing waarmee het vinden van foto's niet meer afhankelijk is van omschrijvingen zou een drastische verandering betekenen. “Je kunt beschrijven wat je concreet op een foto ziet, hoe je de situatie op de foto interpreteert en wat de relatie is tussen de elementen op de foto”, legt Kleppe uit. “Het doel van je onderzoek of archief bepaalt echter hoe je een foto uiteindelijk beschrijft. Daardoor blijft het soms toch lastig om een

foto te vinden terwijl die op zich goed gedocumenteerd is.”

Kies de meest typerende foto

Om de herkenning van foto's verder te onderzoeken, lanceert Kleppe eind februari samen met het Nederlands Fotomuseum de game ‘Mijn Icoonfoto's’. Deelnemers krijgen tijdens dit spel verschillende foto's te zien van bekende gebeurtenissen. De opdracht is om de meest typerende foto te selecteren. Kleppe is benieuwd of hij verschillen tussen generaties kan ontdekken in de manier waarop mensen historische gebeurtenissen herkennen. De game kan online worden gespeeld en vanuit een schoolbank met een *touchscreen*. Deze schoolbank wordt onder andere op stations, in bibliotheken en gemeentehuizen geplaatst. Zo komt het onderzoek op een speelse manier ook naar de mensen toe.

www.martijnkleppe.nl

Dataset, proefschrift en game

Martijn Kleppe werkt als onderzoeker aan de Erasmus Universiteit Rotterdam op de afdeling Geschiedenis van de Erasmus School of History, Culture & Communication (ESHCC).

Kleppe's dataset ‘Foto's in Nederlandse Geschiedenis Schoolboeken (FinGS)’ is beschikbaar via het online archiverings-systeem EASY van DANS. Kleppe hoopt dat de dataset gebruikt wordt door historici voor toekomstig onderzoek of door informatici voor bijvoorbeeld de ontwikkeling van beeldherkenning. De dataset bestaat uit 411 omslagen van geanalyseerde Nederlandse geschiedenis-schoolboeken en 5.344 gecodeerde foto's uit deze schoolboeken.

Het proefschrift ‘**Canonieke Icoonfoto's. De rol van (pers) foto's in de Nederlandse geschiedschrijving**’ verschijnt eind februari 2013 bij uitgeverij Eburon.

De game ‘**Mijn Icoonfoto's**’ kan gespeeld worden via www.mijnicoonfotos.nl.

Oratie Rens Bod: ‘Humanities are the next Big Thing’

Joris van Zundert

Op 14 december accepteerde Rens Bod met zijn oratie getiteld *Het einde van de Geesteswetenschappen 1.0* de nieuwe leerstoel in Digitale en Computationale Geesteswetenschappen aan de Universiteit van Amsterdam. In de hem bekende stijl – Bod wist 7.000 woorden te gebruiken waar 6.000 woorden de conventie is – legde hij uit hoe de scheiding der geesten tussen natuurwetenschappen en geesteswetenschappen goeddeels artificieel is. Het is een scheiding die onder meer gebaseerd is op ideeën van Dilthey, die meende dat er geen patronen te ontdekken waren in de menselijke geschiedenis, maar dat er uitsluitend van interpretatie sprake kon zijn. Dat neemt echter niet weg dat de geesteswetenschappen altijd al een intrinsieke band hebben onderhouden met technologie. Een band die

ver terug gaat, zelfs tot vóór de tijd waarin Roberto Busa een aanvang maakte met de computerondersteunde concordantie op de werken van Thomas van Aquino (1949). Rens Bod stelt dat het juist die technologie is, met name computertechnologie, die beide disciplines opnieuw zeer dicht bij elkaar zal brengen. Informatietechnologie stelt geesteswetenschappers in staat patronen te herkennen in grotere hoeveelheden data dan een menselijke onderzoeker kan lezen. De geesteswetenschappen hebben in de laatste decennia een periode gekend waarin een grote slag in digitalisering is gemaakt, en het is nu aan de onderzoekers om op zoek te gaan naar de fascinerende mathematische patronen die te ontdekken zijn in die digitale ‘fuzzy’, incomplete en complexe informatie gegenereerd door mens en maatschappij. Dat de kennis die dit zal opleveren maatschappelijk en economisch relevant is, staat volgens Bod buiten kijf: bedrijven zoals IBM en Apple beweren immers ‘Humanities are the next Big Thing’.

twitter.com/rensbod

Online zoeken in 8.000.000 krantenpagina's

Miluska Dorgelo

De KB heeft acht miljoen krantenpagina's uit Nederland, Nederlands-Indië en West-Indië (Suriname en de Antillen) doorzoekbaar gemaakt. Kranten zoals de Amigoe di Curaçao, het Bataviaasch Nieuwsblad, De West-Indiër, De Waarheid, het Utrechts volksblad, de Maasbode, het Agrarisch Nieuwsblad, dagblad voor Nederland en vele anderen kranten zijn te vinden op de webdienst kranten.kb.nl. U kunt hier grootschalige, geautomatiseerde zoekacties uitvoeren en selecties maken op titel, datum, algemene artikelen, familie-berichten, advertenties en illustraties met onderschrift. In de speciale krant, als bijlage te vinden bij dit nummer, staan voorbeelden van wetenschappers en onderzoekers die de webdienst al hebben gebruikt. Deze krant is uitgegeven ter gelegenheid van de afronding

van het krantenproject. De webdienst kranten.kb.nl is tot stand gekomen in het kader van het Nationaal Programma Grootschalige Onderzoeksfaciliteiten van NWO/Agentschap NL. Voor de oorlogskranten stelde het programma Erfgoed van de Oorlog van het ministerie van VWS aanvullende financiering ter beschikking.

kranten.kb.nl/



Uitdelen kranten bij livegang website

foto Jos Uljee, KB

Ook in 2013: cursus Data Intelligence 4 Librarians

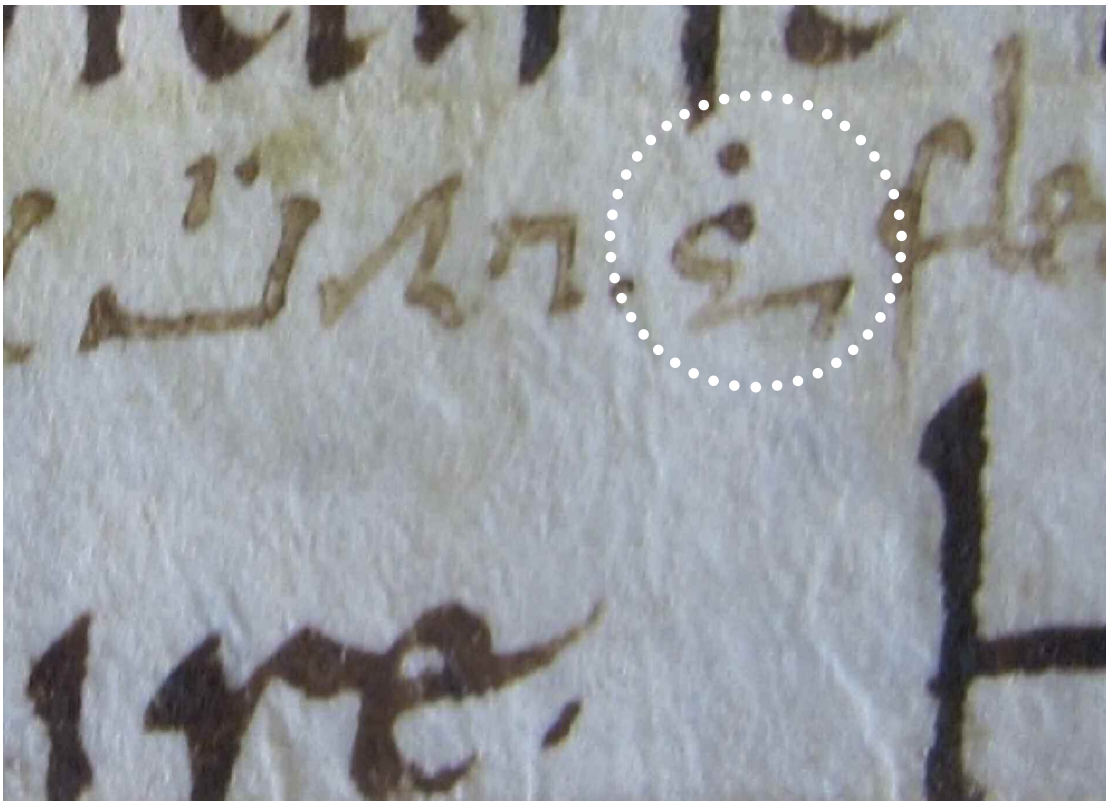
Heidi Berkhout

Sinds het najaar van 2012 bieden DANS en 3TU.Datacentrum de cursus Data Intelligence 4 Librarians aan. De cursus is bestemd voor medewerkers uit alle wetenschappelijke disciplines die meer willen weten over datamanagement en data-archivering. Een deelnemer aan de cursus in 2012: “Deze cursus geeft je een blik op de toekomst. Datamanagement gaat ontegenzeggelijk deel uitmaken van het werk van universiteitsbibliotheken.” Na de cursus kan de cursist onderzoekers op een efficiënte manier adviseren en voorlichten over alle *ins* en *outs* van data. De cursus is een didactische mix van groepsbijeenkomsten, digitaal samenwerken, zelfstudie en het doen van opdrachten. De cursus wordt in 2013 herhaald, meer informatie hierover staat op de website van de cursus. Op deze site staat ook het lesmateriaal, vrij toegankelijk voor iedereen die de cursus (nog) niet volgt.

dataintelligence.3tu.nl/home/



Cursisten aan de slag
foto Carolien van Zuilekom



Een voorbeeld van het Tiroons kortschrift in handschrift. Het omcirkelde teken staat in de *Supertextus notarum Tironiarum* bron Universiteitsbibliotheek Leiden, VLF 48, fol. 2v (© Digital Special Collections)

Weinig geheimen meer voor Tiroons kortschrift

Wetenschappers die worstelen met Tiroonse noten, een antieke en middeleeuwse vorm van stenografisch schrift, beschikken sinds kort over een handige tool.
Evina Steinova

Tiroonse noten zijn aan te treffen in vroegmiddeleeuwse manuscripten. Met uitzondering van een paar specialisten wereldwijd waren ze tot op heden geheimschrift voor wetenschappers. Eén van de redenen is dat het gaat om zo’n grote set tekens. Zelfs met het zeldzame negentiende-eeuwse standaardwerk *Commentarii notarum tironiarum* in de hand, is het heel moeilijk om de complexe grafische symbolen op te lossen tot de woorden waarvoor ze staan. Op papier is het immers moeilijk zoeken omdat de typografische elementen een vaste ordening, zoals de alfabetische, ontberen. Daarbij is het vaak lastig om het ene teken van het andere te onderscheiden, omdat hun vormen minder gestandaardiseerd zijn dan onze schrijffletters.

Zoekpaden
De Supertextus notarum Tironianum: Hypertext-Lexicon der tironischen Noten biedt voor het eerst een hele praktische oplossing: de

symbolen worden gepresenteerd in de vorm van zoekpaden, die vertrekken van basistekens en de gebruiker daaraan dan steeds nieuwe elementen laat toevoegen. De paden reproduceren de beweging van de schrijfhand in het manuscript. Het hypertext lexicon is daarbij nog voorzien van een goede set instructies, en een kort overzicht van de geschiedenis van het Tiroonse kortschrift. De tool is ontwikkeld door Martin Hellmann, auteur van een boek over de aard en het gebruik van Tiroons kortschrift in de Karolingische periode.
www.rzuser.uni-heidelberg.de/~mw8/snt2/n/incipit.htm

GELEZEN

Het gewicht van de auteur. Stylo-metrische auteursherkenning in Middel-nederlandse literatuur.
Joris van Zundert

Onlangs promoveerde Mike Kestemont op *Het gewicht van de auteur*. Stylo-metrische auteursherkenning in Middel-nederlandse literatuur. Met deze studie presenteert Kestemont zich als een wegbereider van de kwantitatieve literatuuranalyse. Het proefschrift geeft een inleiding tot de stylometrie, een methode die minder dan traditionele studie naar auteurschap afhankelijk is van het subjectieve gevoel voor stijl van de literair-historisch expert. Kenmerkend voor de moderne stylometrie is dat er vooral gebruik wordt gemaakt van veel voorkomende en betekenisarme functiewoorden, in plaats van het traditionele gebruik van inhoudswoorden. Vervolgens past Kestemont de methode toe op een casus uit de Middel-nederlandse letterkunde. Een analyse van de rijmwoorden in de *Spiegel historiael* bevestigt eerdere vermoedens over het auteurschap van verschillende tekstdelen. Hoogtepunt van het werk is een onderzoek naar het auteurschap van verschillende onderdelen van de *Lancelot-compilatie*, het belangrijkste handschrift met Arthur-verhalen, bewaard in de KB. Hier blijkt de ingelaste roman *Moriaen* stilistisch diepgaand te verschillen van de rest van de compilatie, en eerder thuis te horen in het gezelschap van *Karel ende Elegast*. Als het proefschrift een gebrek heeft, dan is het dat het misschien soms te voorzichtig is. Dat laat onverlet dat Kestemont met dit proefschrift een jaloersmakend goede studie heeft geschreven, die de meerwaarde van stylometrie voor het vakgebied duidelijk laat zien, en die tot spannende resultaten en nieuwe bewijsvoeringen leidt.
www.mike-kestemont.org

COLUMN

Verslaafd aan e-data

Ik durf het bijna niet op te schrijven, maar ooit – het zal omstreeks 1985 zijn geweest – discussieerde ik met een vriend over de vraag of wij wel of geen gebruik zouden gaan maken van een apparaat dat toen in opkomst was, de zogenoemde pc. Ik was indertijd in het bezit een dure schrijfmachine met een klevend lint dat de laatste zestien tekens die je had getikt van het papier kon trekken. Wat al een enorme verbetering was ten opzichte van de situatie daarvoor, toen je tikfouten verwijderde met Tipp-Ex of met een wit stukje tape. Mijn vriend dacht dat de pc vooral door de generaties na ons gebruikt zou gaan worden. Ik twijfelde nog: misschien was het ook nog iets voor ons. Mijn god, wat waren we – nog geen dertig jaar oud – bedaagde heren! En wat zaten we ernaast! Een jaar later schreef ik mijn docu-

toraalscriptie op de eerste Mac die ik indertijd voor dat doel had gehuurd. De huur was belachelijk hoog, ik meen 500 gulden per maand, maar inmiddels had ik elders ervaren hoe handig zo’n computer (intern geheugen 128 kb) kon zijn.

De mobiele telefoon heb ik bewust een tijd tegengehouden.



foto Leo van Velzen

den. Ik vond het volkomen flauwekul dat je altijd bereikbaar moest zijn. “Als mensen me willen bellen dan wachten ze maar tot ik thuis ben”, zei ik, er totaal aan voorbijgaand dat het voor mijzelf nuttig zou kunnen zijn om onderweg anderen te kunnen bereiken. Mijn eerste mobiele telefoon ben ik heel lang blijven gebruiken (‘Deze doet het nog prima, waarom zou ik overstappen?’) en het heeft ook relatief lang geduurd voordat ik mijn eerste smartphone aanschafte.

Inmiddels ben ik volkomen connected. Ik check mijn email zelfs terwijl ik in de rij voor de kassa sta en ik ga niet meer op vakantie

zonder een tablet. Sterker nog, in het buitenland is een van de eerste vragen in een hotel of restaurant: ‘Hebben jullie wifi?’

En half jaar geleden keek ik nog verbaasd en soms geërgerd naar mensen die in elkaars gezelschap vooral op hun smartphone zaten te kijken. Inmiddels doe ik dat zelf ook. Weliswaar niet de hele tijd, maar af en toe. En de verleiding om te kijken is doorlopend aanwezig. De diagnose is snel gesteld: van een slow adaptor ben ik veranderd in een homo digitalis mobilis, zoals het onlangs in een krant werd genoemd. Verslaafd, afhankelijk, altijd en overal online. Ik vind dat geen positieve ontwik-

keling, maar er is niet echt een weg terug. Matigen zou wel kunnen en daar lees je dan ook steeds vaker over. Controleer slechts drie keer per dag uw email, las minimaal een mediavrije dag per week in, pas op met sociale media want die maken ons in feite sociaal. Het is de boodschap van de zogenoemde Slow Tech-beweging en ik kan hun profeten onmogelijk ongelijk geven. Maar het is makkelijker gezegd dan gedaan. Net als bij andere verslavingen zijn smoesjes snel gevonden. En net als bij veel andere verslavingen is zelfbeheersing, evenwichtig gebruik, uiteindelijk de oplossing. Vanaf nu ben ik – denk ik, hoop ik – iets minder goed bereikbaar. Als het me lukt.

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