



E-DATA & RESEARCH

Jaargang 18 | nummer 1

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, CLARIAH, DANS, KNAW Humanities Cluster, ODISSEI en RKD.

Polifonia verbetert de FAIRness van muzikaal erfgoed

De soundtrack van onze geschiedenis

Van alle culturele ‘producten’ is muziek iets dat mensen samen-brengt in tijd, ruimte en taal. Polifonia is een Europees project van ruim drie jaar dat zich richt op muzikaal erfgoed en een holistische benadering zoekt waarin muziek als fenomeen wordt beschouwd. *Andrea Scharnhorst*

In tien pilotprojecten wordt onderzoek gedaan naar de verzameling, analyse en de interactie met muzikaal erfgoed, en ook het managen van bestaande collecties. Er wordt gewerkt met knowledge graphs gebaseerd op semantic web technologieën die de basis vormen van de verschillende pilotprojecten.

Verborgen patronen
Het doorzoekbaar maken en publiceren van linked data over muzikaal erfgoed is een belangrijk onderdeel van het project. De doelen van Polifonia gaan echter verder dan dat: Verschillende culturele instellingen zijn betrokken bij het project en werken samen aan het verbeteren van de FAIRness van bestaande collecties, die worden verbonden in het Polifonia-portaal. Dit portaal richt zich zowel op een breder publiek dat zich met muziek wil



Credits: Polifonia

bezighouden en verborgen patronen wil zien, als op onderzoekers die op zoek zijn naar gegevensbronnen en nieuwe analysemethoden.

Tools voor muzikaal erfgoed onderzoek
Het Polifonia-portaal geeft toegang tot een open-source toolbox met AI-tools die nieuwe verbanden zullen leggen tussen data over Europees muzikaal en ander cultureel erfgoed. De tools kunnen bijvoorbeeld muzikale verbanden vinden



Credits: Polifonia

tussen Europese liederen. Uit familieverbanden van muzikale elementen in de liederen gezongen op werk of thuis, kan men op migratiestromen in het verleden stuiten. Ook kunnen de tools worden gebruikt om melodieën van verschillende archieven te vergelijken en uit te zoeken hoe artiesten door de eeuwen heen elkaar in hun werk hebben geïnspireerd. Zo kan de Europese culturele verbondenheid worden gereconstrueerd door te bestuderen waar belangrijke muzikale personen, *Vervolg op pagina 3*

INHOUD

- 5 Framing van iconische gebouwen
- 5 Herkennen van phishing en nepwebwinkels
- 6 DANS Data Stations bieden maatwerk voor onderzoekers
- 7



De geuratlas van de Lage Landen is nu beschikbaar

8 Menno Rasch vertelt waarom hij de wetenschap anders wil indelen

E-data & Research liever (tijdelijk) op een ander adres ontvangen? Geef wijzigingen door via edata@dans.knaw.nl.



Een virtuele bibliotheek voor middeleeuwse handschriften

Nieuw handschriftenportaal online

180 middeleeuwse handschriften zijn nu online te bekijken in het recent gelanceerde handschriftenportaal eCodicesNL.

Irene van Renswoude en Mariken Teeuwen
Zeldzame werken van antieke dichters, verluchte getijdenboeken, rechtsteksten in het Oudfries, maar ook beduimelde handschriften met praktische instructies voor een chirurgijn, of recepten tegen haaruitval. Het is een greep uit de handschriften die nu toegankelijk zijn via de website eCodicesNL. Een projectteam van het Huygens Instituut begon drie jaar geleden met de bouw van een nationaal digitaal portaal voor middeleeuwse handschriften in Nederlandse erfgoedinstellingen, met subsidie van het Fonds KNAW Instituten. Doel van het project was om een groot publiek toegang te bieden tot een volledige set digitale afbeeldingen en goede beschrijvingen, via een toegankelijke zoek-interface. Voor het prototype werd samengewerkt met drie instellingen: de Athenaeumbibliotheek in Deventer, Museum Huis van het Boek in Den Haag, en het letterkundig en historische centrum Tresoor in Leeuwarden.

Gebruikersvriendelijk
Een grote uitdaging voor het projectteam lag in

het creëren van een digitale infrastructuur voor een duurzame en interoperabele opslag van de digitale reproducties van de handschriften en hun metadata. De bestaande dataset Medieval Manuscripts in Dutch Collections (MMDC), werd uit de gesloten en verouderde omgeving gehaald en omgezet in een duurzaam XML-formaat, conform een internationaal omarmd TEI-template voor handschriftenbeschrijvingen. Daarnaast werd gekozen voor IIIF. Het team zocht een balans tussen gedetailleerde en vaktechnische informatie en een ontsluiting die meer op een breed publiek gericht is. Het resultaat is een toegankelijke zoekinterface, die de metadata van de gedigitaliseerde manuscripten van de collecties doorzoekbaar maakt.

Digitaliseren op locatie
Bij het inrichten van een workflow voor het creëren en verwerken van hoogwaardige digitale foto's van handschriften werden de kwaliteitsrichtlijnen van Metamorfoze zo goed mogelijk gevolgd. Een team van tenminste twee medewerkers (een fotograaf en een handschriften-deskundige assistent) bezocht de instelling en fotografeerde de handschriften met behulp van de Traveller's Conservation Copy Stand en een hoogwaardige camera. Een groot voordeel van deze werkwijze, waarvan een korte film te zien is op de website van eCodicesNL, is dat de handschriften op locatie kunnen blijven en niet ge-

transporteerd hoeven worden naar een beeldstudio. Er wordt nu gewerkt aan nieuwe subsidieaanvragen die het mogelijk zullen maken om de infrastructuur van het portaal verder uit te bouwen en compatibel te maken met Linked Open Data. De digitalisering van middeleeuwse handschriften wordt intussen voortgezet door de KB/Metamorfoze. Hiermee komt de ambitie dichterbij om alle middeleeuwse handschriftencollecties in Nederland te ontsluiten, zowel voor onderzoekers als voor een breed publiek. <https://ecodices.nl>



Digitaliseren bij Bijzondere Collecties in Nijmegen Credits: Dick van Aalst

National Open Science Festival

Ricarda Braukmann

Op donderdag 31 augustus kwamen meer dan 400 onderzoekers, data stewards, beleidsmakers en andere Open Science enthousiastelingen in Rotterdam samen voor het derde National Open Science Festival om te praten over bestaande en beoogde projecten en initiatieven.

Esther van Rijswijk en Antonio Schettino openden samen met Annelien Bredenoord, Rector Magnificus van de Erasmus Universiteit het event en stipten al verschillende belangrijke onderwerpen aan. Zo werd gesproken over de rol van Open Science in een post-truth world en de klimaatcrisis, en werden projecten gepitcht rondom open hardware en citizen science. Er werd stilgestaan bij het overlijden van Leo Waaijers, een gepassioneerde wetenschapper en idealist, die met zijn 85 jaar nog steeds actief bijdroeg aan de Open Science beweging. Tot slot gaf Hans de Jonge, directeur van Open Science NL, een inkijk in het nieuwe Open Science NL programma.

Na de plenaire sessie verdeelden de deelnemers zich over de community-led breakout sessies. Zoals we gewend zijn van dit festival was er ook dit keer te veel om uit te kiezen: van praktische tools, software en hardware tot aan meer fundamentele discussies over samenwerking en diversiteit en inclusie binnen de Open Science beweging.

Tijdens de pauzes en lunchbreak was er ruimte om met elkaar in gesprek te gaan bij een van de stands van de market place. Na de lunch verdeelden de deelnemers zich nogmaals over twee sets van breakout sessies met onderwerpen zoals reproduceerbaarheid, FAIR data, open software, open educatie en citizen science. Ook een sessie van DANS met praktische tips over het herbruikbaar maken van kwalitatieve data trok veel bezoekers.

Het formele deel van de dag werd afgesloten met een plenaire sessie en een panelgesprek over de ontwikkeling van Open Science NL. Ook werden stellingen bediscussieerd waarin de rol van Open Science in een bredere context werd geplaatst. Er werd gediscussieerd over het combineren van wetenschap en activisme en de rol van Open Science hierin. Veel bezoekers bleven nog lang napraten tijdens de borrel waarin discussies verder werden verdiept en nieuwe samenwerkingen werden bevestigd. Alle presentaties zijn beschikbaar in de Open Science Festival verzameling op Zenodo.

<https://zenodo.org/communities/osf2022nl/>

GEHOORD & BIJGEWOOND



IFLA Credits: RKD

DARIAH Annual Event 2023

Andrea Scharnhorst

Sinds 2017 organiseert DARIAH ERIC jaarlijks een community bijeenkomst. Het DARIAH coördinatieteam van DANS organiseert dit event met trots. Het event is uniek omdat het de deelnemende landen en instellingen, de DARIAH bestuursorganen en werkgroepen een forum biedt om elkaar te ontmoeten. Daarnaast is er ruimte om elkaar te informeren over de voortgang in het onderzoek en de ontwikkeling van de onderzoeksinfrastructuur in het conferentiegedeelte. In de laatste jaren is het conferentiegedeelte steeds groter geworden en heeft het zijn eigen plek ingenomen in de grotere groep van conferenties op het gebied van de Digital Humanities. Verschillende partners worden bij elkaar gebracht zoals musea, bibliotheken, archieven en instellingen die materialen beschikbaar stellen maar ook datacenters en onderzoeksengineers komen samen in het ontwikkelende onderzoeksveld van de Digital Humanities.

Het DARIAH event vond in juni plaats in Hongarije en telde ruim 200 deelnemers uit meer dan 30 landen. Met het onderwerp ‘Cultural Heritage Data as Humanities Research Data?’ was het event gericht op cultureel erfgoed data en werd een nieuwe trend weergegeven in de traditionele samenwerking tussen de Digital Humanities gemeenschap en de GLAM (Galleries, Libraries, Archives and Museums) gemeenschap. De omvang van de conferentie onderstreept de uitgesproken steun van de gemeenschap in het hele DARIAH-netwerk, terwijl het

ook steeds aantrekkelijker wordt voor onderzoekers die nog niet actief zijn in DARIAH, om zo potentiële gebruikers en toekomstige makers van de infrastructuur aan te trekken. Zo trekt is het evenement een springplank voor potentiële gebruikers en makers van de infrastructuur van de toekomst. Nicole Emmenegger en Nils Arlinghaus van het Nederlandse Thematic Digital Competence Centre for the Social Sciences and Humanities (TDCC-SSH) bezochten het DARIAH event: “De conferentie was inspirerend en heeft ons veel nieuwe inzichten gegeven. We hebben veel nieuwe mensen ontmoet en inspiratie gehaald uit de presentaties en discussies. Het was een mooie manier om de gemeenschap te leren kennen!”

Een korte film van het 2023 evenement is beschikbaar op YouTube. <https://www.dariah.eu/>

IFLA Art Library Conference

Dominique Jurgen

Van 21 tot 25 augustus 2023 vond in Rotterdam het IFLA World Library and Information Congress plaats. De IFLA Art Library Section organiseerde samen met de Rijksmuseum Research Library een eigen satellietconferentie op 17 en 18 augustus, met op 19 augustus de mogelijkheid om verschillende bibliotheken in Den Haag (o.a. RKD – Nederlands Instituut voor Kunstgeschiedenis), Amsterdam (o.a. Stedelijk Museum) of Antwerpen (o.a. KMSKA) te bezoeken. Het thema van de conferentie was Big Ideas, Challenging Questions.

Collega's op het gebied van kunstbibliotheken (en aanverwanten) van over de hele wereld kwamen samen om te praten over de toekomst van cultureel erfgoed – en in het bijzonder kunst en visuele informatie. Enkele van de belangrijkste gespreks-onderwerpen waren: Linked Open Data, Collection as Data, samenwerking op Europese en wereldwijde schaal en het aantrekken van jonge mensen naar het werkveld van kunstbibliotheken.

Beide dagen begonnen met inspirerende keynotes. Sally Chambers (bestuurslid van DARIAH, Digital Research Infrastructure for the Arts and Humanities) sprak over ‘The Art of Interconnection: Integrating Art, Visual Cultures and Artistic Practices into Digital Research Infrastructures’. Met de nadruk op internationale (Europese) samenwerking en Collection as Data. Tamar Evangelestia-Dougherty's keynote was getiteld ‘Art Libraries are Everything Everywhere all at Once’. Ze sprak onder andere over de uitdagingen als directeur van de onlangs samengevoegde Smithsonian Libraries and Archives en het belang van kunstbibliotheken als geheel.

Daarnaast was er gelegenheid om het museum te bezoeken en gesprekken te voeren met internationale collega's en boekverkopers. <https://2023.ifla.org>

ESRA 2023

Mara Verheijen

Van 17 tot 21 juli vond het tweetaaljaarlijkse congres van de ESRA (European survey research association) plaats in Milaan. Dit evenement, dat elke twee jaar wordt gehouden, heeft tot doel onderzoekers, beleidsmakers, methodologen en statistici uit heel Europa samen te brengen, om kennis te delen en te verdiepen. Het thema van dit jaar was ‘Survey Onderzoek in Tijden van Crisis: Uitdagingen, Kansen en Nieuwe Richtingen’.

Deze ‘tijden van crisis’, gekenmerkt door gebeurtenissen zoals de COVID-19-pandemie, de oorlog in Oekraïne en klimaatverandering, vormen een grote uitdaging voor onderzoek in de sociale- en gedragswetenschappen. Hoe voer je bijvoorbeeld face-to-face interviews uit als langs de deuren gaan eigenlijk onmogelijk is? Zijn er alternatieve methoden beschikbaar? Veel grote internationale vragenlijsten (bijv. De European Value Survey en de International Social Survey Programme) overwegen om hun data-verzamelmingsmethoden op de lange

termijn te wijzigen naar een online methode voor hun dataverzameling. Het Centerdata LISS-panel werd herhaaldelijk belicht tijdens het congres. Het diende als een voorbeeld van hoe online vragenlijstonderzoek efficiënt kan worden uitgevoerd en welke voordelen dit met zich meebrengt.

Het congres was echter niet beperkt tot discussies over deze ‘push to webstrategie’. Andere onderwerpen die aan bod kwamen waren survey methodologie, mixed-modes en mode effects, data management, respons bevorderende maatregelen, tooling, gamificatie en het gebruik van AI in vragenlijstonderzoek. Voor meer informatie over de besproken sessies, zie website. Over twee jaar zal het congres voor het eerst in Nederland, in Utrecht, worden gehouden.

<https://www.europeansurveyresearch.org/conf2023/prog.php>

CLARIAH Tech Day – 29 juni

Janessa Vleghert

CLARIAH organiseert enkele keren per jaar een Tech Day, waar gebruikers van de CLARIAH-onderzoeksinfrastructuur in gesprek gaan met andere geesteswetenschappers en ontwikkelaars. De Tech Day op 29 juni stond in het teken van data. De vraag hoe de implementatie van nieuwe hulpmiddelen binnen de kunsten en geesteswetenschappen kon worden vergroot, stond hierbij centraal. Tijdens het panel bleek dat vrijwel alle deelnemers van de Tech & Data Day het roerend eens waren dat iedereen profijt zou hebben van een basisniveau digitale geletterdheid. De vraag hoe dat eruit ziet, is een volgend punt. Van welke uitgangswaarde gaat men uit en waar leggen we de verantwoordelijkheid neer? Misschien moeten er twee verplichte MA-cursussen komen die te zijner tijd kunnen worden uitgebreid. Wellicht zijn er ook wel verschillende standaarden voor digitale geletterdheid, behorend bij verschillende sets aan vaardigheden. Er werd volop gebrainstormd.

Het doel van de Tech & Data Day is om de samenwerking tussen disciplines te verbeteren. De meeste ontwikkelaars en gebruikers was het al opgevallen dat er veel data en tools zijn, maar dat onderzoekers niet weten dat deze er zijn of hoe ze gebruikt kunnen worden. De ontwikkelingen gaan razendsnel; het is een veeleisende taak om die bij te houden. Dit soort bezorgdheden spelen gelukkig bij iedereen en het was mooi om op deze interactieve en interessante dag te zien hoe iedereen vanuit zijn eigen perspectief hierover nadenkt en ermee omgaat. <https://www.clariah.nl/events>

Vervolg van pagina 1

Soundtrack Polifonia

zoals componisten, door de jaren heen zijn geweest en hebben opgetreden. Het samenbrengen van zulke verschillende informatiesoorten op een machine-leesbare manier maakt een heel nieuwe kijk op, en een veel beter begrip van, het muzikaal erfgoed in historische context mogelijk. Het project loopt nog tot april 2024.

Europees consortium met Nederlandse expertise

Het Europese project wordt gecoördineerd door Valentina Presutti (Associate Professor aan de Universiteit van Bologna) en er zijn meerdere partners uit Nederland: Beeld & Geluid, het Meertens instituut en DANS. Het ontwikkelen van ontologieën wordt geleid door Albert Meroño Peñuela (hoogleraar computerwetenschappen aan het King's College in Londen) die zijn academische carrière in Nederland begon met een promotietraject aan de Vrije Universiteit (Frank van Harmelen) en de KNAW eHumanities groep. Peñuela heeft een substantiële bijdrage geleverd aan het CLARIAH-project en is momenteel voorzitter van de DARIAH ERIC Working group on Artificial Intelligence and Music. Dit onderstreept - opnieuw - de bijdrage van Nederlandse instellingen en hun interne netwerken aan onderzoeksinnovatie in Digital Humanities.

Kennis makkelijker toegankelijk

Peter van Kranenburg, onderzoeker bij het Meertens instituut leidt twee pilot-projecten. Een van de projecten richt zich op de Orgelencyclopedie (1997-2010) een grote verzameling over Nederlandse Orgelen met een schat aan informatie. In Polifonia wordt de gehele inhoud digitaal beschikbaar gemaakt en in een knowledge graph gepubliceerd. Hierdoor is de inhoud voor iedereen makkelijk en snel digitaal toegankelijk en kan de rijke informatie door machines beter doorzocht en gekoppeld worden aan andere collecties in Europa. De basis van de data vormen ontologieën over muziek die beschrijven hoe de muziek gedocumenteerd en uitgevoerd is, en ervaren wordt.

<https://polifonia-project.eu/>



Credits: Polifonia



Ruben Dood Credits: CBS

Chief Data Officer uit idealisme

Ruben Dood aangesteld als CDO voor het CBS

Met jouw achtergrond is het niet vreemd dat jij bent aangesteld als Chief Data Officer. Waarop ligt bij jou de focus? Op het interne deel, dus alles wat met data binnen de CBS-muren te maken heeft of meer de rol naar buiten?

Mijn rol is vooral naar buiten gericht. Vanuit de maatschappij bestaat een grote vraag naar statistisch onderzoek. Het CBS publiceert al veel statistisch materiaal op Statline en open data. Daarnaast leveren wij als CBS nog Aanvullende Statistische Diensten en bieden wij de mogelijkheid om via Remote-Access onderzoek te doen op onze microdata. Vooral op de laatste twee terreinen heb ik mij de laatste jaren ingezet. Als CDO kan ik daar goed gebruik van maken. Door met externe partijen om de tafel te zitten en bijvoorbeeld mee te denken over hoe om te gaan met de data governance. Dat wil zeggen dat we alles doen om optimaal data te kunnen delen en daarbij niet alleen de veiligheid en betrouwbaarheid, maar ook de ethiek niet uit het oog verliezen. Ook het volgen en implementeren van nieuwe ontwikkelingen speelt daarbij een belangrijke rol.

Over wat voor ontwikkelingen praat je dan? Zit er iets spannends aan te komen?

Ontwikkelingen zoals PPT (Privacy Perserving Techniques), AI, synthetische data en 'digital twins' zijn interessant. Dat is natuurlijk een mooie stap vooruit voor iedereen die met data werkt. Ze maken wellicht breder gebruik van CBS-data mogelijk, maar er moet daarvoor nog veel geregeld en ontwikkeld worden. De eisen qua beveiliging blijven onveranderd streng en die moeten passen binnen de technologische mogelijkheden. Deels vloeien uit deze technieken, mits ze breed worden toegepast binnen Nederland, ook weer

Als "dé" databron in Nederland is het niet vreemd dat het Centraal Bureau voor de Statistiek (CBS) een Chief Data Officer (CDO) heeft aangesteld. Nog minder vreemd is het, dat deze rol is toegekend aan Ruben Dood. Als directeur van onder andere de afdelingen Beleidsstatistiek en Microdat-services is hij als geen ander bekend met het delen van data en alle regelgeving daar omheen.

Hilde van Oirschot en Astrid de Kock

INTERVIEW

manieren voort voor data die gebruikt kunnen worden voor onderzoek. Mijn rol daarbij is vooral om de randvoorwaarden daaromtrent te bewaken.

Als bron van data en kennis is het CBS belangrijk voor andere organisaties. Heeft het CBS zelf nog een belang? Welke toegevoegde waarde kun jij als CDO voor het CBS betekenen?

Het CBS heeft inderdaad veel kennis maar deze is vaak theoretisch. Om hieraan een praktische draai te geven, heeft het CBS an-

dere instanties nodig. Juist die contacten lopen via de CDO. Verder kunnen best practices gemakkelijker worden gedeeld. Dit alles leidt tot maatschappelijke meerwaarde van data. Zo heeft niet alleen het CBS profijt van deze samenwerking maar profiteert de hele maatschappij hiervan mee.

Hoe bedoel je dat laatste?

Overheidsorganisaties, dus ook het CBS, worden betaald met belastinggeld. Dat vind je ook terug in de grote hoeveelheid data die met belastinggeld wordt verzameld. Ik vind dat je als overheid optimaal die verzamelde data in moet zetten, zodat de burger zo veel mogelijk waarde krijgt voor de betaalde belastingen. Dat kan door deze data zo goed en efficiënt mogelijk te gebruiken voor statistisch onderzoek. Onderzoek waaruit conclusies kunnen worden getrokken op basis waarvan beleid kan worden bepaald, uitgevoerd en geëvalueerd. Je moet jezelf een goed beheerder en uitbater tonen van dergelijke voorzieningen. Goed omgaan met de kosten, efficiënt en gebruiksvriendelijk werken. Dat zijn zaken die bovenaan mijn prioriteitenlijstje staan. Ik ben van mening dat belastinggeld optimaal moet worden ingezet. Dat is overigens voor mij ooit de reden geweest om bij de overheid te gaan werken. Eerst bij Rijkswaterstaat en daarna bij het CBS.

Dat is een mooie motivatie! Wil je zelf nog iets kwijt aan de lezers?

Jazeker! Ik vervul mijn rol als directeur én CDO met veel plezier en hoop nu op een productieve en prettige samenwerking met mijn toekomstige gesprekspartners. Mooie data, veilig gedeeld op een dusdanige wijze die 'BV Nederland' ten goede komt!

<https://www.cbs.nl/>

SINDS KORT BESCHIKBAAR

Dit overzicht toont databestanden die recent beschikbaar zijn gekomen bij Centerdata en DANS.

Centerdata

- Behoeftte aan geïnformeerde toestemming in de radiologische gezondheidszorg

In een eerdere versie van E-data (juni 2022) werd beschreven welke door ODISSEI gehonoreerde onderzoeksprojecten een dataverzameling in het LISS panel mochten afnemen. Een van die projecten ging over geïnformeerde toestemming in de radiologische gezondheidszorg. In dit experimenteel onderzoek werden geschreven vignetten vergeleken met videovignetten, waarbij werd gekeken of dit deelnemers hielp de complexe informatie gemakkelijker te verwerken.

lissdata.nl

De data zijn nu ook beschikbaar via LISS Data Archive.
DOI: <https://doi.org/10.17026/dans-26y-5ue8>

Ook sinds kort beschikbaar:

Studies LISS panel

- Centerdata, 2022, LISS panel - Initial Questionnaire – 2022.
DOI: <https://doi.org/10.17026/dans-zbj-2qyt>
- Kesternich; L. de Barys; T. De Schouwer; F. Vermeulen; H. Schumacher; B. Siflinger; maart 2021, Hypothetical Choice Experiment - Pilot study.
DOI: <https://doi.org/10.17026/dans-x9x-5fw5>
- Kesternich; L. de Barys; T. De Schouwer; F. Vermeulen; H. Schumacher; B. Siflinger; juni 2021, Hypothetical Choice Experiment - Main study.
DOI: <https://doi.org/10.17026/dans-zr5-wc3k>
- Bas Denters (University of Twente); Giedo Jansen (University of Twente); Hans Vollaard (Utrecht University); Harmen Binnema (Utrecht University), oktober 2020, Dutch Provincial and Waterboard Election Study 2019.
DOI: <https://doi.org/10.17026/dans-2as-hmaj>
- Evelien Brouwers (Tilburg University, Tranzo), februari 2018, Managers.
DOI: <https://doi.org/10.17026/dans-z4m-2jc7>



Gezondheidszorg Credits: Pixabay

- Evelien Brouwers (Tilburg University, Tranzo), februari 2018, Employees.
DOI: <https://doi.org/10.17026/dans-zbb-xyqe>
- Jonathan Mijs (Erasmus University Rotterdam); Willem de Koster (Erasmus University Rotterdam); Jeroen van der Waal (Erasmus University Rotterdam), september 2020, Social Inequality in the Netherlands: Perceptions, Attitudes and Explanations.
DOI: <https://doi.org/10.17026/dans-x6s-z3ma>
- Centerdata, December 2022 - maart 2023, Politics and Values - Wave 15.
DOI: <https://doi.org/10.17026/dans-xv8-878q>



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

DANS

De volgende datasets zijn open access beschikbaar via de DANS Data Stations of het online archiveringssysteem EASY:
• E.Heunks, 2023, 'De nieuwe

- geomorfogenetische kaart van nijmegen-noord en nabije omgeving',
<https://doi.org/10.17026/AR/SYMI4Z>, DANS Data Station Archaeology, V1
- Dijk, Drs. X.C.C. van, 2023, 'Prehistorische jagers aan de Moerputten',
<https://doi.org/10.17026/AR/V005PI>, DANS Data Station Archaeology, V2
- Pierik, Bob; Van Dam, Petra, 2023, 'Who drinks which water?',
<https://doi.org/10.17026/SS/L8MCKA>, DANS Data Station Social Sciences and Humanities, V2
- Quanjer, B.; Karijo, M.; Kok, J., 2023, 'Municipal conscription height data for the Netherlands 1863-1875 & 1921-1925',
<https://doi.org/10.17026/SS/KMLXUE>, DANS Data Station Social Sciences and Humanities, V2
- Sangers, N.L., 2023, 'Kwantitatieve corpusanalyse stem-elementen in educatieve teksten',
<https://doi.org/10.17026/SS/PRMJRF>, DANS Data Station Social Sciences and Humanities, V1
- Weiss, Laura; Benjamins, Laura, 2023, '10 Years Up',
<https://doi.org/10.17026/SS/Q1GQJO>, DANS Data Station Social Sciences and Humanities, V1

- <https://doi.org/10.17026/SS/Q1GQJO>, DANS Data Station Social Sciences and Humanities, V1
- Young, Jordan, 2023, 'Workshop observations and interview reports of just engagement in Lightness project',
<https://doi.org/10.17026/SS/A4NU3S>, DANS Data Station Social Sciences and Humanities, V2
- Ophuis-Cox, MSc F.H.A. (Open University) (2022): The Effect of Retrieval Practice on Fluently Retrieving Multiplication Facts in an Authentic Elementary School Setting. DANS.
<https://doi.org/10.17026/dans-z8h-c463>
- Hei, Dr. M.S.A. de (The Hague University of Applied Sciences); Audenaerde, Drs. I. (The Hague University of Applied Sciences) (2023): Quick Scan Co-creation in Higher Education. DANS.
<https://doi.org/10.17026/dans-xgc-dagw>
- Maiti, ir. A. (Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente) (2023): UAVPal. DANS.
<https://doi.org/10.17026/dans-z55-6gt4>
- Stokal, M. (Wageningen University & Research); Vriend, P.; Bak, M.P. (Wageningen University & Research); Kroeze, C. (Wageningen University & Research); Wijnen, J. van; Emmerik, T. van (Wageningen University & Research) (2023): Data from: River export of macro- and microplastics to seas by sources worldwide. DANS.
<https://doi.org/10.17026/dans-xaa-kug9>
- Pan, Dr Q. (Erasmus MC) (2023): R code for transcriptomics analysis of mpox virus infection in skin organoids. DANS.
<https://doi.org/10.17026/dans-xzc-46u4>
- van Rijn, Dr. L.J. (Amsterdam UMC) (2023): Data Repository: Configuration TREYESCAN. DANS.
<https://doi.org/10.17026/dans-xes-wdbf>
- Dejaco, B (HAN) (2021): Concurrent validity and reliability of a VR software to assess shoulder range of movement. DANS.
<https://doi.org/10.17026/dans-xdm-q2pc>



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

AGENDA

5 oktober • Rotterdam

Seminar 'Ik wil FAIR handelen - Hoe ga ik om met mijn kwalitatieve data?'
Er wordt veel gesproken over de FAIR-principes van onderzoeksdata. Maar hoe breng je ze in de praktijk voor je kwalitatieve data?
<https://www.surf.nl/agenda/ik-wil-fair-handelen-hoe-ga-ik-om-met-mijn-kwalitatieve-data>

12 oktober • Online

Lezingen - Synthetische data als nieuw onderdeel van het data tijdperk?
Online lezingenmiddag over Synthetische data, om in te gaan op de filosofie en praktische aspecten rond Synthetische data.
<https://www.npsso.net/evenementen/aankomende-evenementen/synthetische-data-als-nieuw-onderdeel-van-het-data-tijdperk>

23-26 oktober • Salzburg

International Data Week
De IDW2023 brengt een gemeenschap samen

van datawetenschappers en data stewards; onderzoekers uit alle domeinen; experts op het gebied van data, interoperabiliteit en informatie uit alle domeinen; industrieleiders, ondernemers en beleidsmakers.
<https://internationaldataweek.org/idw2023/>

23-29 oktober

International Open Access Week
'Community over Commercialization' is het thema voor de Internationale Open Access Week van 2023. Dit thema moedigt een openhartig gesprek aan over welke benaderingen van open wetenschap de belangen van het publiek en de academische gemeenschap voorop stellen en welke niet.
<https://www.openaccessweek.org/>

2 november • Utrecht

ODISSEI Conference 2023
Gericht op de sociaal-wetenschappelijke gemeenschap. Samen discussiëren over data, methodes, infrastructuur, ethiek en theoretisch werk over digitale en computationele methodes in sociaal-wetenschappelijk onderzoek.

<https://odissey-data.nl/nl/2023/09/odissey-conference-for-social-science-in-the-netherlands-2023-2/>

2 november - 8 februari • Online

FAIR-IMPACT
Synchronisation Force Workshops
De workshops zijn bedoeld om gemeenschappelijke uitdagingen en prioriteiten te bespreken met betrekking tot het in de praktijk brengen van de FAIR-principes. Het betreft een reeks van zes virtuele sessies.
<https://fair-impact.eu/events/synchronisation-force-events/synchronisation-force-2nd-workshop-november-2023>

30 november • Utrecht

CLARIAH Annual Conference
De conferentie van dit jaar richt zich op Digital Humanities voor toekomstig onderwijs & onderzoek. Iedereen die actief wil zijn rond dit thema is uitgenodigd, van beginners tot doorgewinterde experts, van taalwetenschappen tot kunstgeschiedenis.
<https://www.clariah.nl/nl/evenementen/clariah-->

annual-conference-2023

8 december • Amersfoort

Symposium 'Oral History, een stap verder'
Voor historische verenigingen, individuele beoefenaars, professionals, vrijwilligers van migrantengroepen, museum- en archiefmedewerkers en iedereen die meer wil leren over oral history.
<https://sprekendegeschiedenis.nl/8-december-2023-symposium-oral-history-eeen-stap-verder/>

24 januari • Amsterdam

OH-SMART Metadata-workshop
Ben je een restaurator of een museummedewerker en maak je gebruik van interviews als onderzoeksmethode in relatie tot collecties? Dan is deze workshop interessant voor jou. Kom meer te weten over het deponeren van collectieinterviews en hoe deze te publiceren en archiveren.
<https://www.uva.nl/en/shared-content/faculteiten/nl/faculteit-der-geesteswetenschappen/evenementen/2024/01/oh-smart-workshop-24-januari-2024.html>

Framing van iconische gebouwen onderzocht

Stadse iconen in beeld in de MediaSuite

Hoe worden iconische gebouwen als de Euromast en het Evoluon in de media geframed? Onderzoeker Vincent Baptist ontdekte dat hij de MediaSuite breder kan inzetten dan hij dacht. *Erica Renckens*

Als oud-student Mediastudies van de Universiteit van Amsterdam was Vincent Baptist al bekend met de functionaliteiten van de MediaSuite. “Ik kende de MediaSuite als tool om systematisch en specifiek te kunnen zoeken in audiovisueel archiefmateriaal en de gevonden beelden te kunnen annoteren”, vertelt de onderzoeker van de Erasmus Universiteit Rotterdam. “Maar in dit project ben ik de tool ook anders gaan gebruiken en dat heeft me aangenaam verrast.”

Verder ontwikkelen

In zijn project FICI (Framing Industrial City Icons) onderzoekt Baptist hoe de Euromast en het Evoluon, als iconische gebouwen in respectievelijk Rotterdam en Eindhoven, in de media in beeld werden gebracht in de tweede helft van de 20e eeuw. Voor dit project ontving

hij een CLARIAH Research Fellowship. Aan de hand van onder andere zijn ervaringen zullen de tools uit CLARIAH’s digitale infrastructuur voor de geesteswetenschappen verder ontwikkeld worden.

Symbolisch gebruik

“Toen ik begon met het annoteren van de beelden, realiseerde ik me al snel dat ik eerst een stap terug moest doen”, vertelt hij. “Ik ben gaan exploreren zoals ik dat ook in een stadsarchief zou doen: bladeren door het materiaal om zo een beter overzicht te krijgen. Dat heeft me geholpen mijn onderzoeksvragen aan te scherpen. Het was alsof de MediaSuite met me meedacht.”

Hierdoor zag hij dat de eerste vermeldingen van de iconische gebouwen meestal plaatsvonden in de context van een gebeurtenis. “Bij de Euromast is dat bijvoorbeeld de Floriade, waarvoor deze in 1960 werd gebouwd. Daarna wordt het gebruik meer symbolisch. Zoals in een reportage over de prostitutieproblemen in Rotterdam begin jaren 80, toen er bij de Euromast en op andere plekken werd geprotesteerd tegen het stadsbestuur.”

Architectuur

Ook ging de onderzoeker zich meer verdiepen in wat een gebouw nu precies iconisch maakt.



De Rotterdamse Euromast in aanbouw (1959), ter gelegenheid van de Floriade in Rotterdam in 1960. Credits: Herbert Behrens / Anefo

Dat bracht hem onder andere bij collega’s van de TU München, die meer gericht zijn op architectuur. “Kun je een gebouw op voorhand als icoon ontwerpen? Dat is nu bijvoorbeeld gebeurd met het Depot van Boijmans van Beuningen. Dan zie je dat zo’n gebouw vooral spectaculair moet zijn, zodat het de aandacht trekt.”

Baptist toonde zijn Duitse collega’s de functionaliteiten van de MediaSuite. “Daarvan

waren ze echt onder de indruk. De MediaSuite is natuurlijk alleen voor Nederlandse audiovisuele data, dus dat kenden ze niet. Ze vonden het heel bijzonder hoe makkelijk al dat archiefmateriaal doorzoekbaar is.”

De komende maanden werkt Baptist verder aan zijn onderzoek. Zijn bevindingen zal hij onder andere delen via het blog van de CLARIAH MediaSuite.

<https://mediasuite.clariah.nl/>

Onderzoek naar digitale vaardigheden

Nederlanders vaak minder digitaal vaardig dan ze denken

Veel Nederlanders zijn minder digitaal vaardig dan ze denken. Een aanzienlijk deel van de Nederlanders heeft namelijk moeite met het herkennen van nepwebwinkels en phishingberichten, terwijl zij denken dat ze die wél zouden herkennen.

Mara Verheijen

Dat blijkt uit onderzoek in het LISS panel voor het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK). In dat onderzoek evalueerde Centerdata de kritische digitale vaardigheden van Nederlanders. Goede digitale vaardigheden zijn fundamenteel voor een goed begrip van en deelname aan de huidige en toekomstige samenleving. Het BZK wilde daarom graag weten hoe het precies staat met de digitale vaardigheden van Nederlanders. Zijn zij bijvoorbeeld goed in het uitvoeren van zoekopdrachten en weten zij wat phishingberichten zijn? En zijn er verschillen te herkennen tussen groepen? Waar eerder onderzoek vooral focuste op de meer functionele vaardigheden (‘knoppen-

kennis’), richt dit onderzoek juist op de meer kritische digitale vaardigheden, die bijdragen aan digitaal bewustzijn. Senior onderzoeker Roxanne van Giesen legt uit: “Het gaat er dan bijvoorbeeld niet om of mensen weten hoe je een e-mail verstuurt, maar wel of ze weten hoe je betrouwbare van minder betrouwbare informatie kunt onderscheiden.” Centerdata bracht voor verschillende doelgroepen in kaart welke kritische vaardigheden mensen missen.

Vragenlijstonderzoek

Als deel van het onderzoek werd een online vragenlijst uitgezet in het representatieve LISS panel. De data van 1.392 respondenten werden geanalyseerd. In de vragenlijst werden de digitale vaardigheden van Nederlanders op een subjectieve manier gemeten. Aan panelleden werd eerst gevraagd wat ze (denken te) kunnen. Om de vaardigheden daarnaast ook op een objectieve manier vast te stellen, werd panelleden gevraagd om verschillende taken uit te voeren. Ze moesten bijvoorbeeld iets opzoeken, een product kopen, nepnieuws en phishingberichten herkennen en aangeven

welke berichten wel of niet gedeeld zouden mogen worden op sociale media. Zo kon het eigen oordeel van respondenten over hun kritische digitale vaardigheden vergeleken worden met hun daadwerkelijke vaardigheden: kunnen mensen net zoveel als ze denken te kunnen?

Digitale vaardigheden van Nederlanders

Het onderzoek leverde verschillende interessante inzichten op. Allereerst laat het zien dat er in een aantal groepen ‘splintervaardigheden’ voorkomt. Daarvan is sprake als iemand goed kan omgaan met de ene digitale context, maar tegen problemen aanloopt in een andere context. Dat lijkt bijvoorbeeld te gebeuren bij jongeren. Zij weten vaak veel over het gebruik van social media en mobiele apps maar zijn beduidend minder goed in het herkennen van nepnieuws.

Ten tweede blijkt het belangrijk om op het gebied van digitale vaardigheden extra aandacht te geven aan lager opgeleiden. Zij lijken namelijk meer moeite te hebben met alle kriti-



Credits: Pixabay

sche digitale vaardigheden.

Ten slotte laat het onderzoek zien dat er op het vlak van digitale vaardigheden twee problematische thema’s zijn: het herkennen van phishing en het herkennen van nepwinkels. Want hoewel veel mensen hier moeite mee hebben, realiseren ze zich dit vaak niet. En dat kan serieuze financiële problemen opleveren. In een wereld waarin het aantal nepwebwinkels en phishingberichten toeneemt, is het dan ook belangrijk om bewustzijn van deze problematiek te vergroten en alle mensen op dit vlak digitaal vaardiger te maken.

Het volledige rapport is te vinden op de website van de Rijksoverheid:

<https://tinyurl.com/bdhe8fmc>

Begin 2024 zal de dataset ook beschikbaar worden gemaakt in het LISS data archive:

<https://www.dataarchive.lissdata.nl/>

Eenvoudig data delen en gevoelige data beschermen

DANS Datadiensten brengen data dichtbij onderzoekers

Deze zomer lanceerde DANS een nieuw, disciplinespecifiek, Data Station voor de Social Sciences & Humanities (SSH). DANS directeur Anja Smit: “Wetenschappelijke disciplines hebben specifieke behoeften voor het bewaren en deelbaar maken van data. Daarom werkt DANS aan diensten op maat per onderzoeksdomein. We willen het hergebruik van onderzoeksdata bevorderen door het onderzoekers zo gemakkelijk mogelijk te maken.”

Ingrid Korver



Anja Smit Credits: Bart van Vliet

Smit vertelt dat het opleveren van het Data Station SSH een belangrijke mijlpaal is in het vierjarige strategische programma ‘Focus op FAIR’ (2021-2025). Onderdeel van dit programma is het realiseren van vier domeinspecifieke Data Stations: Archaeology, Social Sciences & Humanities, Life Sciences en Physical & Technical Sciences. De Data Stations Archaeology en SSH zijn al beschikbaar en de andere twee zullen snel volgen.

Thematische aanpak gewenst

Smit vervolgt: “De ontwikkeling van diensten voor onderzoeksdata staat bij veel organisaties prominent op de agenda in het kader van Open Science. Het gaat dan om weten-

schappelijke instellingen, de Nederlandse overheid, de Europese Commissie en wereldwijde organisaties als UNESCO. Maar denk ook aan het pas opgerichte Regieorgaan OpenScienceNL en de NWO-programma’s voor Lokale en Thematische Digital Competence Centers (DCC’s).”

“Binnen de wetenschap is er een groeiende behoefte aan dataservices en -infrastructuren”, verklaart Smit. “Met tijdelijke financiering wordt hieraan gehoor gegeven. Verbreden van die dienstverlening geschiedt deels langs de lijn van organisaties, zoals universiteiten. Tegelijk wordt ook onderkend dat onderzoeksdomeinen verschillende werkwijzen en behoeften hebben als het gaat om dataservices en -infrastructuren. En dat vraagt juist om een meer thematische invalshoek. De ontwikkeling van de DANS Data Stations, met maatwerk voor onderzoekers binnen onderzoeksdomeinen zoals domeinspecifieke metadata, sluit daar dus goed bij aan.”

Samenwerking is key

Succesvolle nieuwe services moeten in deze fase van ontwikkeling van het datalandschap aan de ene kant goed aansluiten bij wat ver-

schillende disciplines nodig hebben, maar anderzijds verbreden en uitstijgen boven de individuele behoeften. “Dat laatste betekent bijvoorbeeld brede toegang geven tot disciplineoverstijgende collecties datasets op basis van gemeenschappelijke standaarden”, zegt Smit. “Er is in Nederland al veel kennis en ervaring voorhanden maar die is nog versnipperd. In steeds meer organisaties zijn data-repositories, trainingsprogramma’s en andere faciliteiten aanwezig, al is dat in verschillende stadia van ontwikkeling. Bij een groeiend aantal gebruikers is schaalbaarheid van de dienstverlening cruciaal en dat vraagt samenwerking.”

Deze overtuiging is dan ook een belangrijke reden voor DANS om samen met ODISSEI en CLARIAH het thematische DCC-netwerk voor de Social Sciences and Humanities (SSH) te leiden. Daar worden immers communities gevormd en versterkt waarin binnen onderzoeksthema’s de samenwerking gezocht kan worden om de ontwikkeling van datadiensten verder te stroomlijnen.

Data beter vindbaar

DANS zoekt ook de samenwerking in de

vernieuwing van de eigen datadiensten. Bijvoorbeeld als het gaat om vindbaarheid, want datasets zijn nog lang niet altijd gemakkelijk te vinden voor onderzoekers. Om dat te verbeteren, is koppeling tussen datacollecties en verschillende diensten nodig. “Om ervoor te zorgen dat de in DANS gedeponeerde datasets op zo veel mogelijk plekken in de wereld gevonden kunnen worden, zijn metadata van de datasets ook opgenomen in andere zoeksystemen”, vervolgt Smit. “De metadata uit het Data Station SSH zijn bijvoorbeeld ook opgenomen in de Nederlandse Metadata Portal van ODISSEI en DANS is onderdeel van de Europese Data Catalogue van CESSDA.”

Breed loket

De ODISSEI Portal maakt dataverzamelingen van een toenemend aantal repositories vindbaar en kan daarmee uitgroeien tot een breed ‘loket’ voor onderzoekers in de sociale wetenschappen. Het doel van de Portal is dat SSH-onderzoekers data kunnen vinden en toegang kunnen aanvragen via een gezamenlijk portaal zodat je niet meer afzonderlijk de websites van de repositories hoeft te raadplegen. En daarmee is Smit terug bij de DANS Data Stations. De Data Stations Physical & Technical Sciences en Life Sciences zullen de komende periode beschikbaar komen: “Beide nieuwe stations gaan we in nauwe samenwerking met andere organisaties inrichten waarmee we van elkaars expertise gebruik maken en de domeinspecifieke datasets daadwerkelijk dichtbij onderzoekers wordt gebracht”, besluit Smit haar betoog trots.

<https://ssh.datastations.nl/>

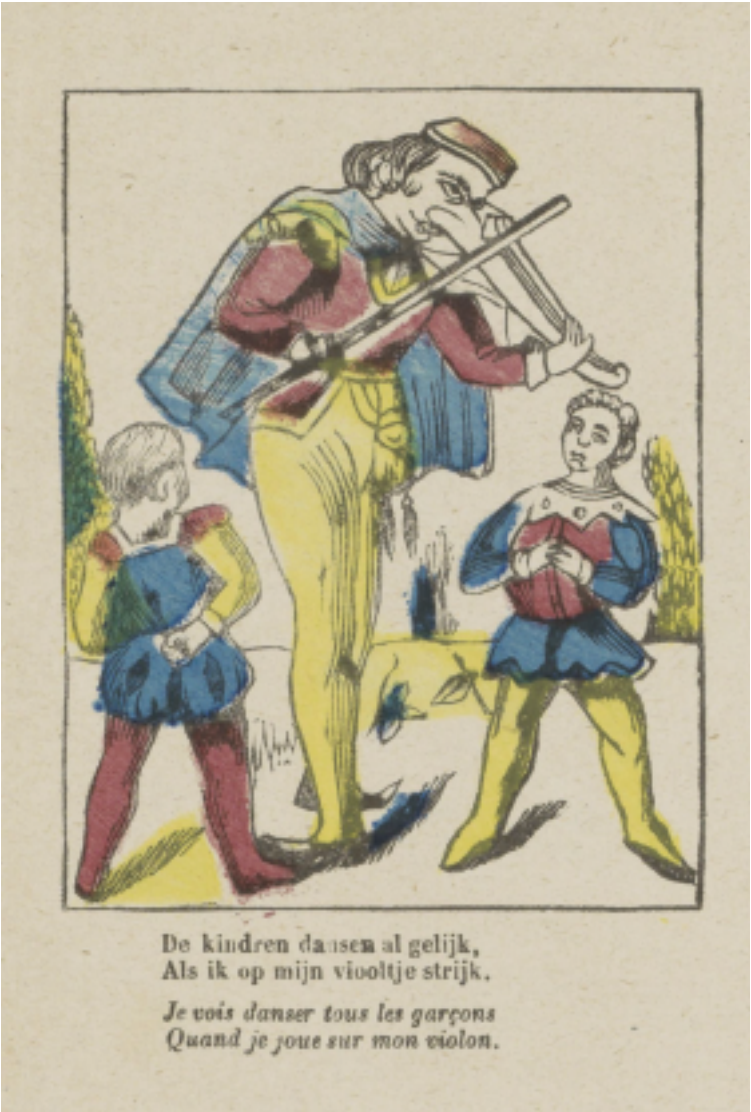
INTERVIEW

‘Gedeponeerde datasets makkelijker vindbaar omdat metadata zijn opgenomen in andere zoeksystemen’

Rijksmuseum:
Spotprent op neuzen,
Franciscus Antonius
Beersmans, 1866-1902

In oktober verschijnt
NeusWijzer. De geur-
atlas van de Lage
Landen. Het boek
brengt inzichten over
geuren uit verschil-
lende wetenschaps-
disciplines samen.

Mathilde Jansen



Digitalisering maakt geuronderzoek in geesteswetenschappen mogelijk

Onderzoek zet schijnwerpers op ons reukorgaan

De NeusWijzer die deze maand verschijnt, is verbonden aan het Europese onderzoeksproject Odeuropa dat onderzoek uitvoert naar het historisch geurerfgoed van Europa. “Het Europese geuronderzoek is de afgelopen jaren ontploft”, vertelt projectleider Inger Leemans. In NeusWijzer worden allerlei wetenschappelijke inzichten samengebracht. “Er zijn een paar grote doorbraken geweest in de biologie en in de chemie over de werking van ons reukorgaan. Sinds covid zijn er bovendien meer mensen die zich realiseren hoe belangrijk het is om te ruiken, niet alleen voor je welzijn maar ook voor je cognitie. En vanuit de geesteswetenschappen zijn er de laatste jaren meer studies verschenen over geur als cultureel verschijnsel.”

Kraskras door literaire bronnen

Hoewel we in Nederland veel pioniers hebben gehad in het geuronderzoek, is er nog maar weinig wetenschappelijk onderzoek gedaan naar het Nederlandse geurlandschap, vertelt Leemans. Frankrijk daarentegen heeft een lange geschiedenis als het gaat om parfumerie, en in Engeland zijn ook allerlei boeken verschenen over dit onderwerp. Vooral binnen de geesteswetenschappen heeft het reukorgaan in

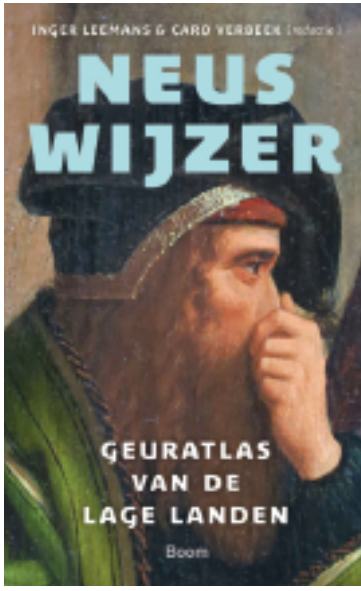
Nederland nooit eerder zo in de belangstelling gestaan als nu. Volgens Leemans heeft dat alles te maken met de digitaliseringsslag die er de afgelopen decennia gemaakt is. “Geur is een multidisciplinair onderwerp. Dus je moet kriskras door literaire bronnen, medische archiefstukken. Te veel om handmatig te doen.” De onderzoeker vertelt hoe ze het hele Woordenboek der Nederlandse Taal heeft doorgespit op geurwoorden, maar dan wel de digitale versie. “Als je dat analoog had moeten doen, was dat een hels karwei geweest.”

Rijk historisch geurlexicon

Vaak wordt beweerd dat het Nederlands in vergelijking met andere talen weinig geurwoorden kent. Maar daar valt volgens Leemans wel wat op af te dingen. “Het onderzoek dat we hebben verricht voor de NeusWijzer geeft aan dat we een rijkere geurtaal hebben dan we altijd dachten. Als je je blik verruimt naar dialecten en regionale talen, zoals we hebben gedaan met een enquête onder 1850 respondenten, kom je verrassend veel geurwoorden tegen.” Maar er zijn ook veel woorden verdwenen. Zo kende het Nederlands vroeger wel 28 verschillende woorden voor ‘muf’. Woorden als ‘duftig’, ‘verdolsemd’, ‘eumig’. Dat die

woorden nu bijna niet meer voorkomen, heeft volgens de onderzoeker alles te maken met technologische ontwikkelingen. “Tegenwoordig hebben we overal verwarming, dus in onze huizen is het veel minder vochtig.” De technologie neemt niet alleen geurtjes weg, maar neemt ook de functie van onze neus over. “Er is apparatuur om geuroverlast te meten, en op alle levensmiddelen staat tegenwoordig een houdbaarheidsdatum. We zijn veel minder afhankelijk geworden van onze neus.”

<https://odeuropa.eu/neuswijzer-de-allereerste-nederlandse-geurwoordengids/>



KORT

Centerdata lid van wetenschappelijk kernteam European Social Survey

De European Social Survey (ESS) is een wetenschappelijk cross-nationaal onderzoek dat sinds 2001 in heel Europa wordt afgenomen. Het brengt attitudes, meningen en gedrag patronen van de bevolking in de verschillende Europese landen in kaart. Elke twee jaar worden face-to-face interviews gehouden op basis van een nieuwe, representatieve steekproef.

Software voor online dataverzameling

Centerdata heeft veel ervaring in het ontwikkelen van gespecialiseerde software en tooling voor gegevensverzameling voor internationale onderzoeken. Deze ervaring zal van groot belang zijn voor de overstap die ESS nu wil gaan maken van face-to-face naar self-completion. Hiervoor worden zowel online als papieren vragenlijsten ingezet. Centerdata zal een platform ontwikkelen dat de online dataverzameling voor alle ESS-landen centraliseert. Zo worden de efficiëntie van het dataverzamelingsproces en de datakwaliteit verbeterd. Deze tooling zal ook de implementatie van gecentraliseerde papieren self-completion vragenlijsten van de ESS faciliteren. (Mara Verheijen, Centerdata)

Administratieve data als voorbeeld

Het belang van PIDs voor het citeren van datasets

Dit artikel laat zien hoe het ODISSEI Portal het citeren van CBS microdatasets mogelijk maakt door aan deze sets Persistent Identifiers (denk bijvoorbeeld aan DOI's) toe te kennen. Een DOI maakt het eenvoudig om snel datasets te kunnen vinden.

Angelica Maineri en Thomas Groen

Het gebruik van persistent identifiers (ook wel PIDs) is een van de steunpilaren van de FAIR (Findability, Accessibility, Interoperability and Reusability) principes. Een PID helpt bij het identificeren van een specifieke bron, net als een burgerservicenummer (BSN) dat doet om een specifiek individu in Nederland te identificeren. Het bekendste voorbeeld van een PID is een DOI, een unieke reeks van cijfers en letters die een specifiek 'object' op het internet identificeert, zoals een gepubliceerd artikel of een dataset. Bij het citeren van datasets in academische publicaties zou het gebruikelijk moeten zijn om PIDs te gebruiken, zoals gespecificeerd door de Joint Declaration of Data Citation Principles (JDDCP).

Voordelen van citeren van data
Het citeren van data in wetenschappelijke pu-



Credits: Cottonbro studio via Pexels

blicaties heeft meerdere voordelen. Ten eerste verhoogt het de betrouwbaarheid van een studie omdat dezelfde data kan worden gebruikt om de resultaten te reproduceren. Daarnaast maakt het databronnen zichtbaarder en eenvoudiger te vinden voor hergebruik door andere academici binnen hetzelfde onderzoeksveld. En het is ook een manier om erkenning te geven aan het werk van degenen die de data hebben verzameld en verwerkt. De meeste data repositories (zoals Dataverse,

Zenodo en Figshare) voorzien gebruikers van richtlijnen hoe er geciteerd kan worden maar geven ook een aantal daadwerkelijke referenties.

Citeren van administratieve data
Het ODISSEI Portal maakt het eenvoudig om data uit de uitgebreide CBS microdata-catalogus te citeren via een aanbevolen datareferentie. Elke CBS-registratie in het ODISSEI Portal verwijst naar een data-ontwerp, name-

lijk een verzameling datasets die dezelfde structuur delen (bijvoorbeeld variabelen en methodologie) maar betrekking hebben op verschillende perioden (bijvoorbeeld jaren). De titels en afgekorte titels van de CBS datasets en data-ontwerpen zijn echter niet altijd consistent, wat het moeilijk kan maken om een specifieke databron te vinden. Hier komt een DOI van pas, zoals beschreven in het onderstaande voorbeeld.
Gebruikers van de dataset KINDOUDERTAB kunnen ontdekken dat het data-ontwerp soms wordt aangeduid als 'In de Gemeentelijke Basisadministratie Persoonsgegevens (GBA) ingeschreven personen en hun juridische ouders' of gewoon als 'Kindoudertab: personen en hun juridische ouders'. Dankzij het ODISSEI Portal kan de bron nu worden geciteerd met behulp van de DOI 10.57934/0b01e410801f9401. Een hyperlink brengt de lezers naar de juiste webpagina waar KINDOUDERTAB op het ODISSEI Portal wordt beschreven, ongeacht welke naam er wordt gebruikt. Door middel van DOI's stelt het ODISSEI Portal onderzoekers in staat eenvoudig te verwijzen naar datasets. Dit biedt voordelen doordat het hen in staat stelt om de data die ze gebruiken in hun CBS-projecten gemakkelijk te citeren, waardoor hun onderzoek transparanter en betrouwbaarder wordt. Bovendien kunnen lezers gemakkelijk toegang krijgen tot uitgebreide informatie, zoals methodologie, over de geciteerde dataset.
<https://portal.odissei.nl>

COLUMN

Kantel de wetenschap

Wetenschap delen we door- gaans in met Alpha-, Bèta- en Gammawetenschappen. Alpha wetenschappers bestuderen de producten van de menselijke geest, Beta de natuur en natuur- lijke wetmatigheden en Gamma het menselijk handelen. Elk deel is weer onderverdeeld in discipli- nes zoals taalkunde, biologie of psychologie en daarna in nog spe- cifiekere vakgebieden. Deze in- deling is grotendeels bepalend in welke omgeving wetenschappers hun werk doen. Scheikundigen zitten op een gang met collega scheikundigen, sociologen gaan naar congressen met andere soci- ologen, filosofen zitten in een re- dactie van een tijdschrift met filo- sofen, en informatici dienen een proposal in met andere informa- tici. Natuurlijk zijn er vele dwars- verbanden, overlappende vakge- bieden en interdisciplinaire sa- menwerking, maar primair is het wetenschappelijk onderzoek zo

georganiseerd. Niet zonder rede- nen natuurlijk, maar dit heeft ook een keerzijde.

W ant het is ook denkbaar dat je wetenschap anders in- deelt. Je kan bijvoorbeeld kijken hoe het onderzoek wordt uitge- voerd. Dan valt vooral op hoe groot de diversiteit is binnen een discipline. Er zijn wiskundigen die voor een krijtbord staan en wiskundigen die hun werk niet zonder een supercomputer kunnen doen. Er zijn biologen die gedrag



Menno Rasch

van dieren observeren, maar ook die geavanceerde instrumenten gebruiken in een lab. Er zijn eco- nomen die werken met big data, en die kwalitatief onderzoek doen. Een manier om zo'n indeling te maken kan door naar de rol van data in het onderzoek te kijken. Hoe wordt de data verzameld? Hoe gaat het analyseren van de data in zijn werk, en hoe werkt het interpreteren van de analyses? Welke uitdagingen zijn er bij het omgaan met de data? Zijn die er bij de hoeveelheid van de data, of bij de complexe statistische analy- ses, of bij de betrouwbaarheid? Door onderzoek met soortgelijke uitdagingen bij elkaar te brengen, kan er een andere indeling van de wetenschap ontstaan. Of, om het in organisatietermen te zeggen: we kantelen de wetenschap, of: we gaan naar een matrixorganisa- tie. Zo'n kanteling brengt uiteenlo- pende wetenschapsgebieden bij

elkaar. Wetenschappers die elkaar nu nooit tegenkomen, kunnen sa- menwerken met wetenschappers die op een vergelijkbare manier met data omgaan. De taalkundige die miljoenen tweets analyseert ontmoet de historicus die het ge- digitaliseerde VOC-archief analy- seert. De antropoloog die vooral interviews afneemt gebruikt de- zelfde software en technieken als de geneeskundige die patiënten bevraagt. De natuurkundige die enorme hoeveelheden data 'real time' moet opvangen, deelt best practices met de econoom die be- talingsverkeerdata gebruikt.
V oor het bouwen van een zo generiek mogelijke infra- structuur kan dit een alternatieve benadering zijn. Het realiseren van infrastructuur voor alle onder- zoekers uit een vakgebied is vaak een hele lastige opgave, juist van- wege die grote diversiteit aan on- derzoeksmethoden. Maar infra-

structuur voor het analyseren van big data, of infrastructuur voor het vastleggen van kwantitatieve data, voor het toepassen van ma- chine learning op grote hoeveel- heden tekst – daar zijn grote groe- pen onderzoekers mee te onder- steunen.
Hoe krijgen we echter deze we- tenschappers bij elkaar aan tafel? Ze komen elkaar niet tegen op de gang, bij congressen of in be- stuurlijke organen. Maar er is hoop: dit blad brengt lezers uit vele verschillende vakgebieden samen rondom data. Zien jullie kansen?

Menno Rasch
Menno Rasch is directeur van de afdeling Digitale Infrastructuur van het Humanities Cluster bij de KNAW. Hij geeft de column door aan Anja Smit.