

De ontwikkeling van een risicotaxatietool voor provincies

Criminaliteit bestrijden met een datagedreven aanpak

Witwaspraktijken, mensenhandel en productie van drugs: ze komen regelmatig voor op bedrijventerreinen. De provincies Gelderland en Noord-Brabant wilden graag in kaart brengen welke terreinen het meest kwetsbaar zijn voor deze vormen van ondermijning.

Patricia Prüfer en Seyit Höcük

Avans Hogeschool en Centerdata gebruikten daarvoor een datagedreven aanpak: samen ontwikkelden ze een risicotaxatietool die gemeenten houvast moet gaan bieden bij het bestrijden van criminaliteit.

Op sommige bedrijventerreinen is het een komen en gaan van bedrijven en op andere zijn relatief weinig aansluitingen op het energienetwerk. Wat betekent dat? En kun je deze informatie koppelen aan andere data om zo beter te voorspellen waar ernstige criminaliteit plaatsvindt? Dergelijke vragen stonden centraal in dit project in opdracht van de provincies Gelderland en Noord-Brabant.

Waar vindt ondermijning plaats?

In samenwerking met Avans Hogeschool verzamelde Centerdata een jaar lang data over bijna duizend bedrijventerreinen. Eerst werden daarvoor mogelijke indicatoren geïdentificeerd op basis van literatuuronderzoek en gesprekken met criminologen en andere experts. Het ging daarbij bijvoorbeeld om leegstand, overlastmeldingen, verhuizingen, KvK-inschrijvingen en energiegebruik. Over deze indicatoren werden gegevens verzameld uit allerlei bronnen, zoals de basisregistratie van het Kadaster en informatie van het CBS en netbeheerders.

Met data science technieken gingen de onderzoekers vervolgens op zoek naar verbanden in die goudmijn aan data. Met behulp van machine learning ontwikkelden ze een risicotaxatietool



Credits: Pixabay

om te identificeren welke bedrijventerreinen mogelijk kwetsbaarder zijn voor ondermijnende activiteiten.

Het onderzoek wijst bijvoorbeeld uit dat het risico op drugsgelateerde criminaliteit groter is op verouderde bedrijventerreinen waar veel autobedrijven zitten. Terreinen waar veel bedrijven ontbonden worden en waar veel kapitaal geplaatst wordt op aandelen van naamloze vennootschappen, zijn juist weer kwetsbaarder voor witwaspraktijken. Wie met de risicotaxatietool op zoek gaat naar deze ondermijnende activiteiten, zal dus geattendeerd worden op bedrijventerreinen met een dergelijke ‘giftige cocktail’ van kenmerken.

Inzet van de risicotaxatietool

Op 30 mei 2022 werd de risicotaxatietool officieel overgedragen aan de provincies. De tool wordt nu beschikbaar gesteld aan gemeenten, zodat zij zicht krijgen op welke bedrijventerreinen meer of minder kwetsbaar zijn voor onder-

mijning. Dit zal gemeenten helpen deze ontworpende vorm van criminaliteit te voorkomen en bestrijden.

Vervolgstappen

Dit project laat al zien hoe data science handvatten kan bieden voor beleid, in dit geval om de veiligheid in gemeenten te verbeteren. Er zijn verschillende interessante vervolgstappen. Zo kan deze kant-en-klare tool in principe worden uitgebreid naar alle Nederlandse provincies. Door vergelijkbare data te verzamelen bij andere provincies en deze op te nemen in de risicotaxatietool kunnen ook daar kwetsbare bedrijventerreinen worden aangewezen in een overzichtelijk dashboard.

Daarnaast kan er op basis van verschillende soorten data gekeken worden naar andere vormen van ondermijnende criminaliteit, zoals illegale arbeidsuitbuiting en faillissementsfraude. Ook kan er mogelijk verder worden gekeken dan alleen bedrijventerreinen.

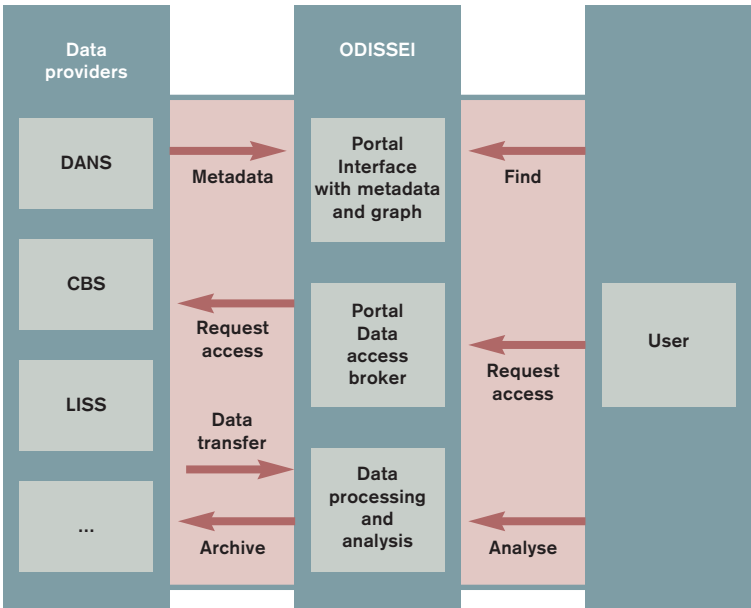
Metadata van meerdere providers zoeken, vinden en aan elkaar knopen

Hey Google, Hey Siri, Hey ODISSEI !

Of we binnenkort met de woorden ‘Hey ODISSEI’ gegevens over datasets (metadata) kunnen oproepen, valt nog te bezien maar een zoekmachine voor metadata uit de sociale wetenschappen is in de maak. Astrid de Kock

ODISSEI (Open Data Infrastructure for Social Science and Economic Innovations) is een samenwerkingsverband van verschillende universiteiten en organisaties zoals CBS, DANS, SURF en de VU, om statistisch onderzoek binnen – met name – de sociale wetenschappen te verbeteren en vergemakkelijken. Vanuit deze onderzoeksinfrastructuur is een

portal in ontwikkeling dat het uitvoeren van wetenschappelijk onderzoek niet alleen makkelijker, maar ook leuker maakt. Om metadata van verschillende instellingen te vinden en aan elkaar te knopen, wordt gebruikgemaakt van een linked data model en is er gekeken naar internationale standaarden voor metadata zoals DDI (Data Documentation Initiative) en de metadata standaard van CESSDA (Consortium for Social Science Data Archives). De metadata van verschillende data-providers wordt in het portal op dezelfde manier opgesteld waardoor deze behalve goed doorzoekbaar,



Schematische weergave van het ODISSEI Portal

Vervolg op pagina 2

E-DATA & RESEARCH

Jaargang 17 | 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 het Rijksmuseum.

INHOUD

3
Kranten uit de 17e eeuw doorzoeken

4
Deze datasets zijn sinds kort beschikbaar

5


Oorlogsbrieven als digitale dataset

6
Nieuwe CLARIAH-PI blikt vooruit

7
Unieke geluidsopnamen te beluisteren

8
Gastcolumnist Tom de Smet over een stoffige werkplek en (linked) open data

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

Scan deze QR-code met een smartphone om de website van E-data te bezoeken. [edata.nl](mailto:edata@dans.knaw.nl)

Open Science Festival

Ricarda Braukmann

Op 1 september organiseerde het Nationaal Programma Open Science (NPOS) voor de tweede keer het Nederlandse Open Science Festival.

Het uitverkochte event trok 300 bezoekers naar de VU in Amsterdam waaronder onderzoekers, service providers en beleidsmakers. Wie zijn ticket niet op tijd had besteld kon het plenaire gedeelte via de live stream volgen.

Vanaf half 10 stroomden de bezoekers binnen om vervolgens in een van zeven pre-festival workshops de dag te beginnen. Deelnemers konden discussiëren over de ontwikkelingen rondom pre-prints en pre-registratie van onderzoek, Diamond Open Access, of een sessie bijwonen over de EOSC en FAIR data. Ook het gebruik van synthetische data om gevoelige data toegankelijk te maken werd besproken.

Tijdens de lunch konden mensen netwerken en bij een van de vele stands praten met de Nederlandse service providers. Naast DANS waren bijvoorbeeld ook het eScience Center en de Open Science Community aanwezig.

Na de lunch werd het Festival officieel geopend door Jeroen Geurts (Rector VU) en Hilde van Wijnngaarden (Directeur Universiteitsbibliotheek VU) waarna het publiek zich weer kon verdelen over diverse parallele sessies. Deelnemers konden leren over software management en programmeer skills, discussies bijwonen over de implementatie van FAIR, citizen sciences, en community building. Ook lanceerde ODISSEI de ODISSEI Portal in een van de parallele workshops (zie pagina 1).

Het hoogtepunt van het festival was de uitreiking van de Open Science Awards door Robbert Dijkgraaf (Minister van Onderwijs, Cultuur en Wetenschap). Vijf projecten die zich kenmerken door

GEHOORD & BIJGEWOOND



Credits: Pim Rusch

hun maatschappelijke relevantie en een directe samenwerking van de onderzoeksdoelgroep werden met deze prijs geëerd. Een van de projecten richt zich op een historische databank van Curaçao en Suriname en de betrokkenen van het onderzoek werden via een live verbinding naar de zaal in Amsterdam gehaald.

Na afloop werd er nog veel genotwerkt tijdens de borrel die een ge-

zellige afsluiting vormde van een geslaagde dag waarin een gevoel van gemeenschap steeds sterk aanwezig was.

<https://opensciencefestival.nl/>

Terugblik op het 25-jarig jubileum van Centerdata

Lilly van Geest
Op 9 juni vierde Centerdata haar 25-jarig jubileum in de aula van Tilburg University. Samen met

partners, opdrachtgevers, oud-medewerkers en collega-onderzoekers proostten we op 25 jaar Centerdata. Het werd een mooie en inspirerende dag waarop we met samenwerkingspartners, opdrachtgevers en oud-medewerkers een beeld schetsten van ons gezamenlijke onderzoek. Dat herinnerde ons er weer eens aan waarom we zo graag doen wat we doen: met dataverzameling en -

analyse kunnen we echt een verschil maken voor mens en maatschappij!

Na zijn officiële opening gaf directeur Marcel Das het woord aan ‘founding father’ Arie Kapteyn. Hij deelde zijn persoonlijke verhaal over de oprichting van Centerdata en benadrukte het belang van probability-based panels. Daarna gaven Lonneke van de Poll-Franse, Marike Knoef, Lian Smits, Patricia Prüfer en Arjan Vliegheart met hun lezingen en gesprekken een mooi beeld van het werk dat Centerdata samen met partners doet.

De sprekers gaven inkijsjes in hun werk en lieten tegelijk zien hoe waardevol de data van Centerdata daarbij zijn. Hoe helpen data de kwaliteit van leven te verhogen? Kan een jeugdzorginstelling met behulp van data sturen op betere ontwikkeling van kinderen? Zijn er betere pensioenkeuzes te maken met data? Wat vertellen data ons over verstandig lenen?

Dank aan alle sprekers en aan iedereen die aanwezig was op deze mooie dag. Om af te sluiten met de woorden van dagvoorzitter Joop Vos: “Het leven is mooi als je met data bezig bent!”.



Wandverslag van 25-jarig jubileum Centerdata Credits: <https://www.creatiefverslag.nl/>

Hey Google, Hey Siri, Hey ODISSEI !

Vervolg van pagina 1

ook vergelijkbaar zijn. Ook wordt het gemakkelijk om 'sleutels' te vinden zodat snel duidelijk wordt via welke variabele bestanden gelinkt kunnen worden.

Snel en efficiënt

De metadata van de bestanden die het CBS ter beschikking stelt via een remote access omgeving worden in het ODISSEI Portal opgenomen. Het doorspitten van honderden pdf-

documenten in de online catalogus is daarmee verleden tijd. Ook metadata van het LISS panel en duizenden sociaalwetenschappelijke datasets die bij DANS zijn opgeslagen zijn vindbaar via de portal. Maar dat is nog maar het begin want ODISSEI ambieert om ook metadata van andere dataproviders zoals NTR, DataverseNL en HSN via het portal vindbaar te maken. En daar blijft het niet bij. Heb je als (potentiële) gebruiker van data gevonden wat je zocht, dan kun je direct via de portal

je toegangsaanvraag bij de desbetreffende instelling indienen. Uiteraard vindt het uitvoeren van het onderzoek in een beveiligde omgeving (zoals de microdata-omgeving bij het CBS) plaats. Een sprong voorwaarts dus, maar niet alleen voor de gebruikers van data. Ook de dataproviders komen aan hun trekken. De leveranciers van de data krijgen eenvoudig inzicht in het ‘wie-wat-waarom’. Wie is geïnteresseerd in welke data en voor welk doel. Het portal stelt metadata beschikbaar

voor internationale dataportalen van CESSDA en de European Open Science Cloud (EOSC).

De eerste resultaten

De afgelopen maanden is via een testportal om feedback van onderzoekers gevraagd. Aan de hand daarvan worden nu verbeteringen doorgevoerd. Zo is bijvoorbeeld de user interface verbeterd en worden links gemaakt naar gerelateerde datasets in de CBS databank StatLine. Op het Open Science Festival

is onlangs de eerste versie van het portal gelanceerd en de komende twee jaar zal ODISSEI verder werken aan deze ODISSEI-zoekmachine die voor eenieder beschikbaar is. Een 'nieuw tijdperk' waarin we via enkele 'klikken' de juiste bronnen, data en expertise kunnen vinden.

Wil je meer weten? Neem dan contact op met info@odissei-data.nl

<https://portal.odissei.nl/>

COLOFON Uitgever: E-data & Research. **Redactieadres:** Anna van Saksenlaan 51, 2593 HW Den Haag, 070-3494450, edata@dans.knaw.nl, edata.nl, @eDataResearch. **Hoofd-/eindredacteur:** Ingrid Korver. **Redactie:** Ricarda Braukmann, Mathilde Jansen, Astrid de Kock, Hilde van Oirschot, Erica Renckens, Mara Verheijen, Suze Zijlstra. **Redactiesecretariaat:** Ingrid van der Elst. **Aan dit nummer werkten mee:** Lilly van Geest, Valentijn Gilissen, Seyit Höcük, Cees Hof, Patricia Prüfer, Tom de Smet, Maaike Verburg, Jerry de Vries, Samantha Willemsen. **Opmaak:** Colette Sloots, Haarlem. **Productie:** Amsterdam University Press. **Druk:** Ten Brink, Meppel. **Oplage:** 4800 papier, 4800 digitaal. **ISSN:** 1872-0374. We hebben getracht alle belanghebbenden met betrekking tot het gebruikte beeldmateriaal te benaderen. Degenen die menen recht te kunnen doen gelden, kunnen zich tot ons wenden. E-data & Research kosteloos ontvangen? Laat het weten via edata.nl. E-data & Research liever (tijdelijk) op een ander adres ontvangen? Geef wijzigingen aan ons door via edata@dans.knaw.nl.

OVERNEMEN ARTIKELEN

Wilt u een artikel uit dit blad overnemen?

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

Zeventiende-eeuwse kranten nu nog beter doorzoekbaar

‘Verrijkte data maken rijker onderzoek’

Het Couranten Corpus bevat zeventiende-eeuwse kranten die op Delpher beschikbaar zijn. Door toevoeging van allerlei metadata zijn ze nu beter doorzoekbaar voor onderzoekers. *Mathilde Jansen*

De kranten waren al toegankelijk via Delpher, maar systematisch onderzoek was lastig doordat de metadata onvolledig en vervuild waren. Bij metadata moet je denken aan data die extra informatie geven over het krantenbericht, zoals het genre: binnenlands nieuws, buitenlandse nieuws, een advertentie of het laatste nieuws. Ook ontbrak taalkundige verrijking, waardoor je bijvoorbeeld door te zoeken op ‘lopen’ alle vormen en spelvarianten van het werkwoord ineens vindt. Onder leiding van taalkundige Nicoline van der Sijs is het zeventiende-eeuwse corpus nu opgeschoond en verrijkt.

Hondjes die zoek zijn
Zo’n driehonderd vrijwilligers werkten mee; zij hebben alle artikelen handmatig overgetikt. Dat was best een uitdaging, vertelt Van der Sijs, vanwege het gotische schrift. “Het zijn vaak kleine letters die

dicht op elkaar staan. Het papier was ook vrij kwetsbaar en er is niet altijd even voorzichtig mee omgesprongen, dus sommige pagina’s zijn zwart of gebold en moeilijk leesbaar. Met name hoofdletters zijn lastig, vooral als het om plaatsnamen of personen gaat die je niet kent.” Toch was de meerderheid van de vrijwilligers erg gemotiveerd, zegt de onderzoeker, omdat het materiaal zo boeiend is. “Je hebt enerzijds het nieuws over de tachtig- en dertigjarige oorlog, wat er gebeurde met de soldaten, maar anderzijds is er ook heel lokaal nieuws. Over hondjes die zoek zijn, mensen die dood gevonden worden in de gracht, diefstal, noem maar op.”

Nederlandse Dataprijs
Om geen fouten te maken werden alle transcripties ook nog eens nagelopen en gecorrigeerd. In 2020 was deze klus geklaard en ontvin-

gen de betrokken partijen (Instituut voor de Nederlandse Taal, Meertens Instituut en de Koninklijke Bibliotheek) de Nederlandse Dataprijs voor dit project. Daarna zijn de metadata bewerkt en is door het INT automatisch de taalkundige verrijking toegevoegd. Nu de teksten online staan is Van der Sijs alweer bezig met het vervolgproject. Er is namelijk nog een aantal extra zeventiende-eeuwse kranten gedigitaliseerd en dan zijn er de kranten uit de 18e eeuw. Ook die zijn deels nog in het gotisch schrift, maar gelukkig kan de computer daar steeds beter mee overweg. “We gaan nu een experiment doen met Transkribus, een computerprogramma dat teksten transcribeert, zodat de vrijwilligers alleen nog maar hoeven te corrigeren. Dat blijft wel nodig, want door gebrek aan context maakt de computer soms fouten die mensen er moeite-loos uithalen. Op deze manier zorgen we voor betrouwbare data en daarmee voor betere onderzoeksresultaten.”

<https://ivdnt.org/corpora-lexica/courantencorpus/>

Grootschalige Online Experimenten

Innovatieve dataverzameling voor sociaalwetenschappelijk onderzoek

Het internet biedt mogelijkheden om grootschalige experimenten online uit te voeren, maar het bouwen van een digitale omgeving hiervoor is ingewikkeld. Het Mass Online Experiment Lab wil dit voor onderzoekers toegankelijker maken. *Suze Zijlstra*

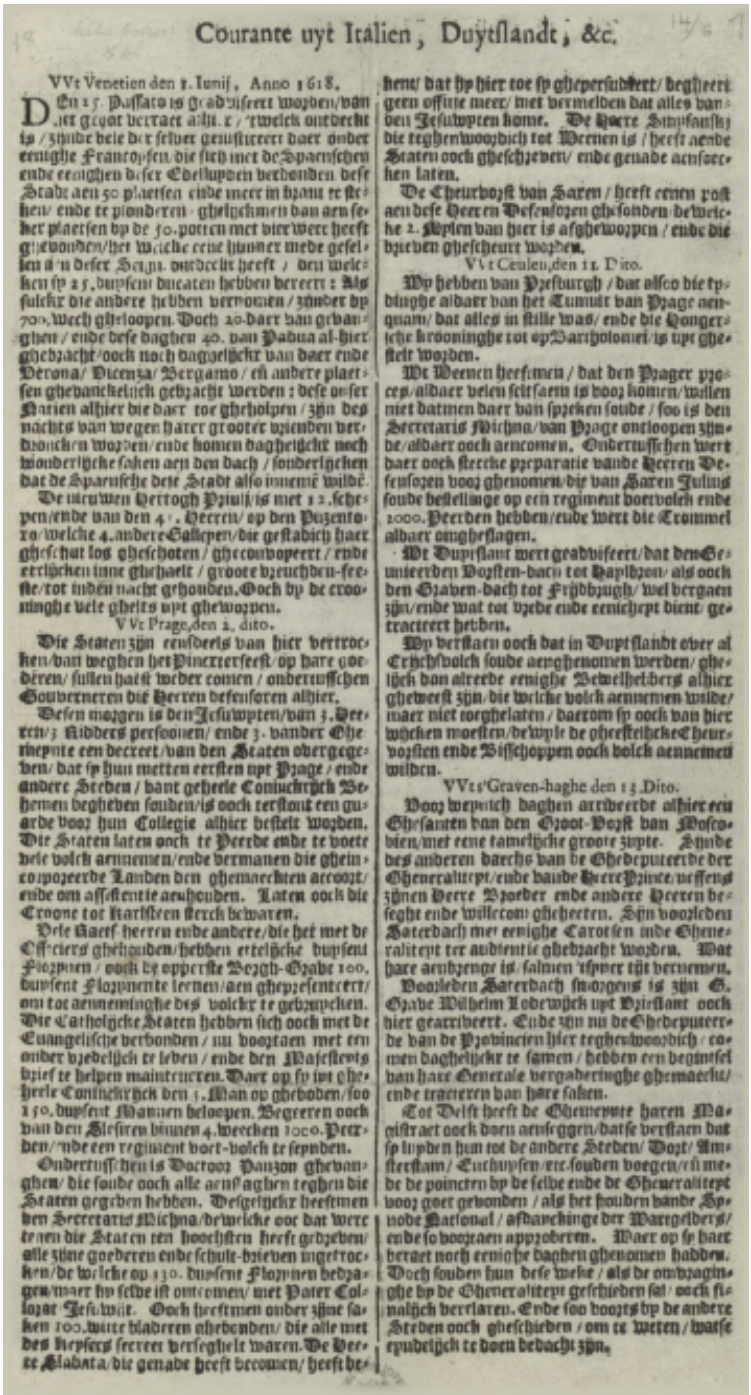
ODISSEI, de nationale infrastructuur voor sociale wetenschappen, ondersteunt op verschillende manieren de ontwikkeling van innovatieve dataverzameling. Een belangrijke pilot die onlangs is afgerond, is het ‘Mass Online Experiment Lab’. Een van de projectleiders van de pilot, Rense Corten (Universiteit Utrecht), licht toe: “Grootschalige experimenten zijn belangrijk voor sociologie - dat gaat immers over grootschalige sociale verschijnselen. Veel kan je in een klein lab goed onderzoeken. Maar voor sommige dingen heb je schaal nodig om de complexiteit van sociale processen te vangen.” Om die schaal te berei-

ken, ontwikkelt het project een virtuele omgeving om met veel participanten tegelijkertijd experimenten uit te voeren. “Dit heeft technische uitdagingen: je wil niet dat het platform hapert. Maar het heeft ook logistieke uitdagingen: je wil bijvoorbeeld dat iedereen tegelijk op een bepaald tijdstip inlogt.”

Efficiënte ontwikkeling
Het ontwikkelen van zo’n omgeving is belangrijk, want veel sociale wetenschappers hebben niet de kennis en programmeervaardigheden om dit zelf te doen. Bovendien is het duur en inefficiënt om voor elk nieuw onderzoeksproject een geheel

nieuw platform te ontwikkelen. Het doel van deze pilot is dan ook geweest om grootschalige online experimenten toegankelijker te maken. In de pilot zijn twee experimenten uitgevoerd gericht op de werking van sociale netwerken: één ging over het verspreiden van desinformatie en de ander over het verspreiden van ziektes. Hiervoor zijn online experimenten gebouwd. “Dit riep verschillende vragen op: welke software is geschikt, hoeveel kan je hergebruiken voor andere experimenten? Je wil een soort gereedschapskist bouwen met bouwstenen die voor andere experimenten hergebruikt kunnen worden.”

Gereedschapskist
Uit de pilot bleek dat deze gereedschapskist in de toekomst waarschijnlijk niet door individuele onderzoekers kan worden ingezet voor nieuwe experimenten. Uiteindelijk zijn er nog steeds programmeurs



De allereerste Nederlandse krant is de Courante uyt Italien, Duytslandt, &c en verscheen op 14 juni 1618. Credits: www.delpher.nl

KORT

ODISSEI Introductie- presentaties

De nationale infrastructuur voor sociale wetenschappen, ODISSEI, biedt onderzoekers veel mogelijkheden. ODISSEI ondersteunt onderzoekers bij ontwikkeling van innovatief computationeel onderzoek, voorbereiding van grantaanvragen (van NWO of Europese grants zoals ERC) en financiert data-toegang via beurzen en kortingen. Nog niet alle onderzoekers die werken bij organisaties die aan ODISSEI deelnemen zijn bekend met alle mogelijkheden. Leden van het ODISSEI Coördinatieteam komen daarom graag langs om hierover te vertellen. Zowel (afdelingen van) deelnemende organisaties, als organisaties die meer willen weten over lidmaatschap, kunnen een presentatie aanvragen. (SZ)
<https://odissei-data.nl/nl/vraag-een-odissei-presentatie-aan/>

SINDS KORT BESCHIKBAAR

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

CBS

Het CBS stelt bestanden beschikbaar voor statistisch onderzoek. Per juni van dit jaar zijn zes Woonbase bestanden beschikbaar gekomen. Woonbase is ingericht om de vraag "Wie woont waar met wie?" te beantwoorden. Hiertoe zijn per jaar drie type bestanden ingericht: een sleutelbestand, kenmerkenbestanden en populatiebestanden. Ook nieuw is het bestand Zorgwoonruimtetab. Zorgwoonruimten zijn verblijfsobjecten waar institutionele huishoudens woonachtig zijn die zorg nodig hebben.



Een volledig overzicht van beschikbare bestanden staat in de microcatalogus op de CBS-website:
cbs.nl/microdata-catalogus



Ook een link naar de bestanden die de afgelopen periode beschikbaar zijn gekomen staat online:
cbs.nl/microdata-overzicht



Credits: iStock

DOI: <https://doi.org/10.17026/dans-zft-buhd>
• Metz, A., oktober 2020, Privacy Perception.
DOI: <https://doi.org/10.17026/dans-xqf-kpru>
• Mu, T., april 2021, Validation of the Individual Ambidexterity Scale.
DOI: <https://doi.org/10.17026/dans-z56-qgzx>
• Okat, D., augustus 2020, Payment apps - Trust in Financial Technology 2020.
DOI: <https://doi.org/10.17026/dans-zqg-4khm>
• Simonse, O., juli 2021, Poor benefits: The role of financial stress in the non-take-up of social welfare.
DOI: <https://doi.org/10.17026/dans-xxu-u8sv>
• Reeskens, T., Halman, L., Luijkx, R., september 2017 - januari 2018, European Values Study 2017 - Main measurement 2017.
DOI: <https://doi.org/10.17026/dans-26g-mv64>
• Rooduijn, M., maart 2021, Experiment populism and anti-immigration views.
DOI: <https://doi.org/10.17026/dans-xsp-ea3v>
• Van der Plas, A., september 2020, About palliative care.
DOI: <https://doi.org/10.17026/dans-z8n-n7pe>
• Van de Ven, N., december 2020, Better than average 2020.
DOI: <https://doi.org/10.17026/dans-28j-ymvg>
• Vromans, R., april 2021, Communicating personalized health information.
DOI: <https://doi.org/10.17026/dans-22s-6gdj>



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

DANS

Sinds kort beschikbaar:
De volgende datasets zijn open access beschikbaar via het online archiverings-systeem van DANS:
• Inspectie Overheidsinformatie en Erfgoed, 2022, 'Archeologische rapporten; vermist en gevonden'.
DANS: <https://doi.org/10.17026/AR/4GVJOK>, DANS Data Station Archaeology, V1
• Sociaal en Cultureel Planbureau (SCP) (2021): Vrouwen in besluitvorming 2020 - VIB2020.
DANS: <https://doi.org/10.17026/dans-xe5-9p9e>
• Verkuyten, prof M. (Utrecht University) (2021): Tolerance in the Netherlands.
DANS: <https://doi.org/10.17026/dans-z4p-zcud>
• Farace, Dr. D. (GreyNet International) (2022): Digital Publishing and Grey Literature.
DANS: <https://doi.org/10.17026/dans-zhs->

fewk
• Deijl, L.A. van der (University of Amsterdam) (2022): A New Language for the Natural Light: Translating the New Philosophy in the Dutch Early Enlightenment (1640-1720).
DANS: <https://doi.org/10.17026/dans-zhq-58ag>
• Raut-Samplonius, Dr. S.K. Ms. (RuG) (2019): Multilateral trade disputes until 2016.
DANS: <https://doi.org/10.17026/dans-x6q-eag3>
• Reckien, Dr.rer.nat. D. (Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, The Netherlands) et al (2022): Plan quality characteristics of Local Climate Adaptation Plans in Europe.
DANS: <https://doi.org/10.17026/dans-xd6-w7pc>
• Ceccon, D. (WagelIndicator Foundation); Medas, G. (WagelIndicator Foundation) (2021): WagelIndicator Collective Agreements Database Dataset with Full Texts and Selected Clauses.
DANS: <https://doi.org/10.17026/dans-zg9-wtij>
• Pan, Dr Q. (Erasmus University Medical Center) (2022): Recapitulating infection, thermal sensitivity and antiviral treatment of seasonal coronaviruses in human airway organoids.
DANS: <https://doi.org/10.17026/dans-xsq-ps55>
• Muijsenbergh, prof. dr. M.E.T.C. van den (Radboud University); Loenen, dr. T. van (Radboud University) (2022): Dakloosheid en Corona.
DANS: <https://doi.org/10.17026/dans-2c7-wksz>
• Zuizewind, Drs. C. (Vilans); Bruinsma, Dr. J. (Maastricht University) (2022): Op weg naar implementatie van de zorgstandaard voor jonge mensen met een dementie.
DANS: <https://doi.org/10.17026/dans-z4j-tmez>
• Latul, MD YP (Amsterdam UMC, location AMC) (2022): Non-invasive biomechanical assessment of the prolapsed vaginal wall.
DANS: <https://doi.org/10.17026/dans-z62-mnyp>
• Meer, M. van der (Wageningen University & Research) (2022): Data from: Same strategies – different categories: An explorative card-sort study of plant-based proteins comparing omnivores, flexitarians, vegetarians and vegans.
DANS: <https://doi.org/10.17026/dans-z9m-rz39>



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

AGENDA

24 - 30 oktober
Open Access Week
Het thema 'Open For Climate Justice' zoekt de verbinding en samenwerking tussen de klimaatbeweging en de internationale open gemeenschap.
<http://openaccessweek.org/>

31 oktober • Den Haag
NPSO lezingenmiddag
Het thema is 'Sensordata en Surveys'.
<https://www.npsso.net/Evenementen/Aankomende-evenementen/sensordata-en-surveys-lezingenmiddag>

3 november • Utrecht
ODISSEI
Conference for Social Science
Voor sociale wetenschappers en data wetenschappers met interesse in computationele sociale wetenschappen.
<https://odissei-data.nl/en/conference-2022>

24 november • Utrecht
CLARIAH & NDE 2022
Het thema is 'Erfgoed ontmoet Digitale Humanities'.
<https://www.eventbrite.com/e/clariah-nde-annual-conference-2022-tickets-416530051597>

29 november • Utrecht
FAIR Data Day
Dit live-evenement draait rond het centrale thema: Het hergebruik van data om de wetenschap vooruit te helpen, het tonen van inspirerende voorbeelden en het samenbrengen van gemeenschappen.
<https://www.aanmelder.nl/fairdataday>

29 november • Utrecht
RDNL Dutch Data Prize
Op 29 november 2022 wordt de Nederlandse Dataprijs uitgereikt tijdens de FAIR Data Day | Reusing Data to Advance Science.
<https://www.aanmelder.nl/fairdataday/dutchdatapriz>

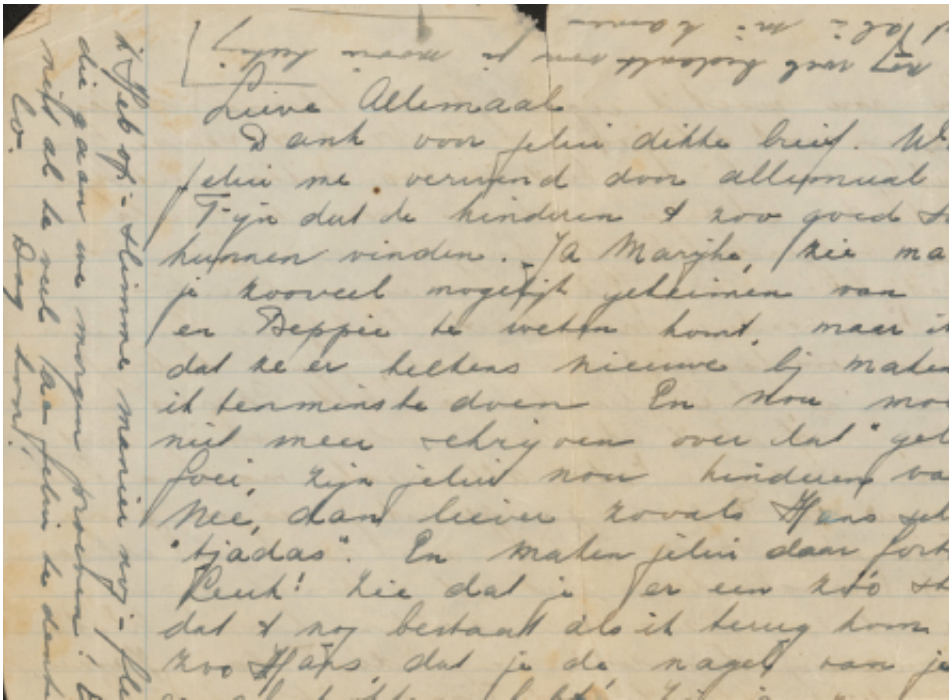
Nieuwe stakeholders gezocht

E-data & Research is het blad over data en onderzoek in de alfa- en gammawetenschappen. We schrijven over recente ontwikkelingen, onderzoeken en bevindingen in dataverzameling en onderzoek. We zijn er voor onderzoekers en dataprofessionals en iedereen die zich betrokken voelt bij de ontwikkelingen in opslag en (her)gebruik van digitale onderzoeksgegevens. E-data & Research verschijnt in oktober, februari en juni per post. Onze artikelen zijn ook via www.edata.nl beschikbaar en het is mogelijk een mail te ontvangen met een link naar de nieuwste E-data.

Wilt u meeschrijven aan E-data?
Nieuwe stakeholders zijn altijd welkom! Als stakeholder mag u meebeslissen over en meeschrijven met de artikelen die in E-data worden gepubliceerd. Hierdoor wordt ook uw netwerk en wetenschappelijk bereik uitgebreid.
Bent u geïnteresseerd in de mogelijkheden als stakeholder? Neem dan contact met ons op via edata@dans.knaw.nl



Beeldmateriaal van het project ‘Oorlog uit Eerste Hand’
Bron: NIOD



Fragment van een brief met tekst in drie verschillende richtingen geschreven
Bron: NIOD Collectie 247, inventarisnummer 1298, scan 36

Oorlog uit Eerste Hand: Oorlogsbrieven van het NIOD (1935-1950) digitaal

Handgeschreven ‘egodocumenten’ automatisch transcriberen

Het NIOD digitaliseert haar bijzondere collectie handgeschreven brieven uit de periode voor, tijdens en na de Duitse bezetting van Nederland en de rekolonisatie-oorlog in Indonesië.

Milan van Lange, Annelies van Nispen en Carlijn Keijzer

‘Oorlog uit Eerste Hand’ (2020-2023) conserveert en digitaliseert “Collectie 247”, een uitzonderlijk deel van het NIOD-archief, bestaande uit persoonlijke correspondentie in tijden van geweld, bezetting, oorlog en vervolging. De ca. 150.000 oorlogsbrieven, kattenbelletjes, ansichtkaarten en enveloppen zijn gescand en worden getranscribeerd en geannoteerd. Dit om een betere (online) vindbaarheid, toegankelijkheid en ruimere bruikbaarheid voor historici en geïnteresseerden mogelijk te maken. In plaats van de scans integraal te laten transcriberen door vrijwilli-

gers, tot voor kort gangbaar, slaat het project een andere weg in.

Van transcripties naar computermodellen

Allereerst is er handgeschreven tekst van ongeveer 1000 scans handmatig getranscribeerd. Deze ‘ground truth’ is vervolgens gebruikt om een computermodel voor automatische handschrijfterkenning te trainen. Hiervoor zijn de functies van READ-COOP’s Transkribus gebruikt. Deze software is voortgekomen uit Europese samenwerkingsprojecten om Handwritten Text Recognition (HTR) voor historisch archiefmateriaal te faciliteren. Het trainingstraject leverde uiteindelijk een computermodel op voor handgeschreven Nederlands uit het midden van de twintigste eeuw, waarbij gemiddeld minder dan 5% van de tekens fout gelezen wordt (Character Error Rate (CER) <5%). Het model wordt gebruikt om de rest van de collectie automatisch te transcriberen. Het genereren en toepassen van het computermodel dat de handgeschreven teksten leest bleek een iteratief proces met veel tussentijds

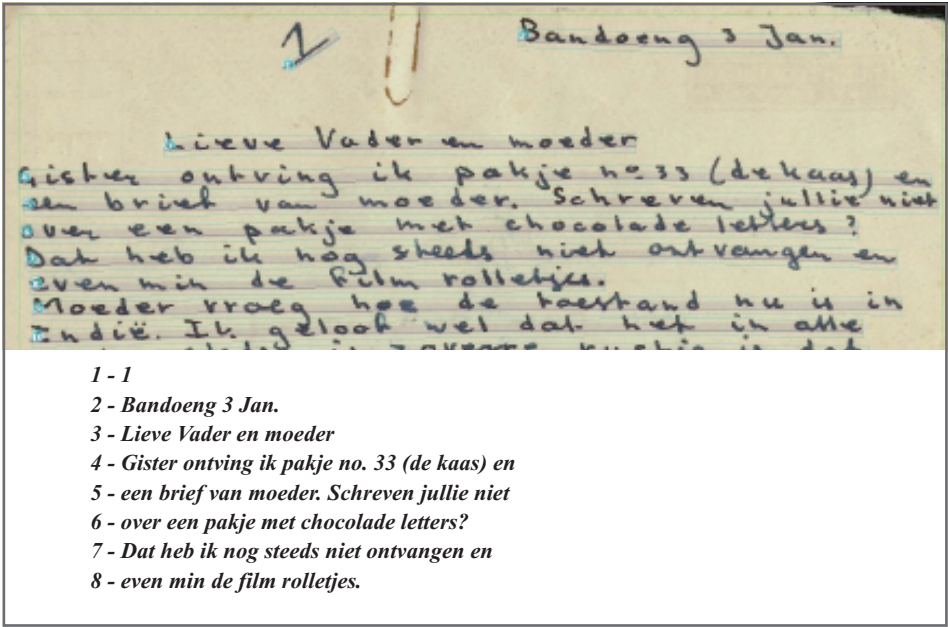
bijsturen. De omstandigheden waarin veel brieven tot stand gekomen zijn en het feit dat er veel verschillende brieven-schrijvers zijn, maken de collectie soms lastig leesbaar voor mens en machine. In tijden van paperschaarste schreven mensen vaak in de rondte of op de kop om toch die laatste mededeling of succeswens nog kwijt te kunnen op die ene schaarse briefkaart. Dergelijke complexe lay-out heeft soms een negatieve invloed op de kwaliteit van de resultaten van automatisch transcriberen. De zeer problematische ‘moeilijke gevallen’ met slechte transcripten gaan we met behulp van crowd sourcing corrigeren. Uiteindelijk worden de transcripten – indien juridisch mogelijk – online vindbaar. Het HTR-model voor handgeschreven tekst uit de periode 1935-1950 wordt via Transkribus en GitHub vrij beschikbaar gesteld aan erfgoedinstellingen en onderzoekers voor (her)gebruik.

Oorlogsbrieven als digitale dataset

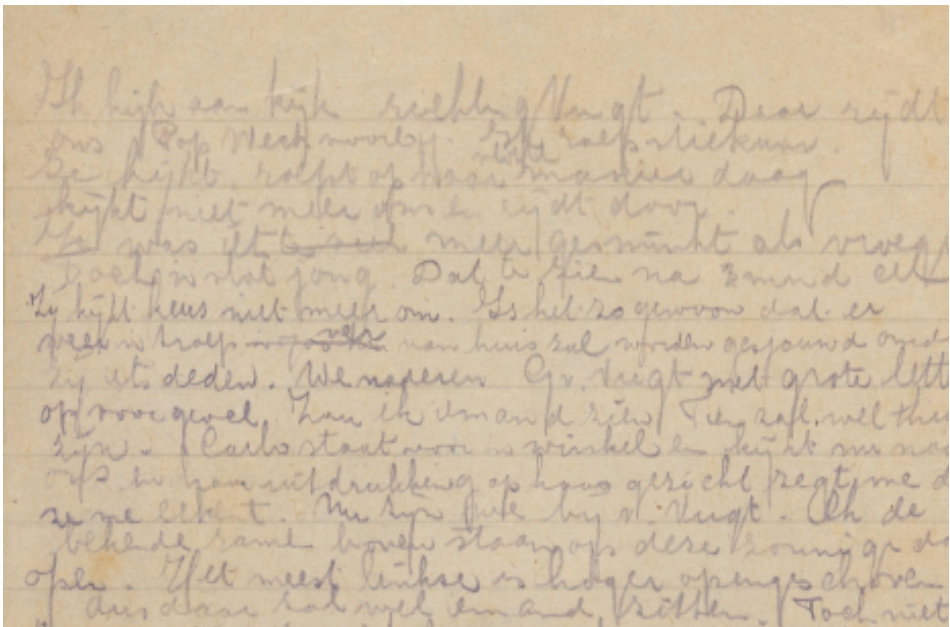
Het transcriberen wordt komend jaar voltooid en er wordt, met automatische classificatie

en crowd sourcing, metadata toegevoegd over personen, plaatsen en organisaties. Dit vormt de basis voor een gestructureerde dataset en opent nieuwe mogelijkheden voor de toepassing van computationele zoek- en analysemethoden. ‘Oorlog uit Eerste Hand’ maakt zo een veelzijdige en bijzondere collectie ‘egodocumenten’ geschikt voor een meer systematische methodologische aanpak in de geschiedwetenschap. Dit kan een belangrijke impuls vormen voor meer schaalbaar onderzoek naar persoonlijke ervaringen van voor, tijdens en na de Duitse bezetting van Nederland en de rekolonisatie-oorlog in Indonesië.

‘Oorlog uit Eerste Hand’ wordt mede mogelijk gemaakt door financiering van het Ministerie van Volksgezondheid, Welzijn en Sport en het Mondriaan Fonds. Meer informatie is te vinden op de website van het NIOD: <https://www.niod.nl/nl/projecten/oorlog-uit-eerste-hand> en in diverse blogs, bijvoorbeeld: <https://www.niod.nl/nl/blog/over-digitalisering-keuzes>



Fragment van een getranscribeerde briefwisseling in de werkomgeving van Transkribus
Bron: NIOD Collectie 247, inventarisnummer 1206



Fragment van een lastig te transcriberen brief voor zowel mens als machine
Bron: NIOD Collectie 247, inventarisnummer 1421, scan 24

Van den Bosch geeft CLARIAH-stokje door aan Van Miert

‘Eerst de winkel goed op orde’

CLARIAH-hoofdonderzoeker Antal van den Bosch heeft zijn taken overgedragen aan Dirk van Miert. Samen geven zij een terug- en een vooruitblik op CLARIAH. Erica Renckens

“We moeten de winkel goed op orde hebben voordat onderzoekers kunnen binnenwandelen. Inmiddels zijn we zo ver dat de deur openstaat”, zo verbeeldt Dirk van Miert de huidige stand van zaken bij CLARIAH, de digitale infrastructuur voor de geesteswetenschappen. Afgelopen juni nam de kersverse directeur van het Huygens Instituut het stokje over van Antal van den Bosch als Principal Investigator (PI) van CLARIAH.

Van den Bosch was als directeur van het Meertens Instituut sinds september 2020 hoofdonderzoeker van CLARIAH. In de tijd dat hij het programma trok, stond de ontwikkeling van de brede technische infrastructuur centraal, zo blikt hij terug.

“In eerste instantie lag de nadruk van CLARIAH op taalkunde, mediastudies en sociaal-economische geschiedenis. Dat kwam zo op natuurlijke wijze voort uit de ontwikkelingen binnen de digitale geesteswetenschappen in Nederland. Maar er was binnen CLARIAH nog weinig overlap tussen deze vakgebieden en bovendien is CLARIAH uiteindelijk bedoeld voor de volledige breedte van de geesteswetenschappen. Daar hebben we de afgelopen jaren samen met de beoogde gebruikers en de verschillende werkpakketten aan gewerkt. Nu breekt de volgende fase aan, waarin we de infrastructuur aan de gebruikers geven. Die roep klinkt ook steeds duidelijker uit het veld.”

Een wereld te winnen

Van Miert zal deze nieuwe stap met een frisse blik begeleiden. “Als literatuur- en cultuur-

historicus ben ik ongetraind in de digitale geesteswetenschappen”, vertelt hij. “Hoewel ik groot voorvechter ben van big data-analyses, laat ik het technische aspect van het onderzoek doorgaans over aan mijn postdocs en aio’s. Ik heb dus inhoudelijk nog veel te leren, en ik zie om me heen dat ik daarin niet alleen ben. Ik schat dat ongeveer de helft van mijn collega’s geen idee heeft wat CLARIAH is en zich ook niet goed kan voorstellen wat er allemaal mee mogelijk is. Daar valt dus nog een wereld te winnen.”

CLARIAH gaat de boer op

Daarvoor moet CLARIAH een proactieve houding aannemen, stelt hij. “We kunnen niet wachten tot mensen vanzelf onze winkel binnenlopen, want die lopen ze voorbij, simpelweg omdat ze niet weten wat we precies

‘We geven de infrastructuur nu aan de gebruiker’

verkopen. We gaan CLARIAH daarom de komende tijd actief naar de onderzoekers brengen, onder andere met trainingen en laagdrempelige toegangspunten. En ook naar de docenten, zodat zij de onderzoekers van de toekomst kunnen lesgeven in de digitale geesteswetenschappen. We moeten hen laten zien en uitleggen hoe dingen werken. Veel mensen worstelen met datasets. De CLARIAH-portal Ineo zal daar bij helpen, maar ook daar is training voor nodig. Naast vertellen, moeten we ook goed blijven luisteren naar wat de geesteswetenschappers nodig hebben. Daar gaan we de komende anderhalf jaar mee aan de slag.”

www.clariah.nl



Antal van den Bosch (l.) en de nieuwe CLARIAH-PI Dirk van Miert
Credits: Simone Wolff

Landschapsanalyse van de Nederlandse erfgoedwetenschappen

Cultureel erfgoed in kaart gebracht

DANS heeft voor het E-RIHS.nl consortium een inventarisatie uitgevoerd van de huidige en beoogde partners en hun relatie met data, infrastructuren en de data-wensen voor het E-RIHS.nl netwerk in oprichting. Maaike Verburg, Jerry de Vries en Cees Hof

De Nederlandse hub van het European Research Infrastructure for Heritage Science (E-RIHS) consortium, E-RIHS.nl, is momenteel in oprichting. Het doel van E-RIHS.nl is om wetenschapsmethoden, onderzoeksfaciliteiten, expertise en data uit het roerende erfgoed, monumenten en archeologie beter toegankelijk te maken. Om meer informatie te verzamelen over de

huidige stand van zaken bij de (potentiële) partners van E-RIHS.nl heeft DANS het project ‘Samenwerken met en aan data: DIGILAB en FAIR’ uitgevoerd. Door middel van interviews en een enquête is een groot aantal (potentiële) partners uit het veld van de erfgoedwetenschappen benaderd om de huidige stand van zaken en hun wensen voor de toekomst met betrekking tot het vinden, delen, archiveren en hergebruiken van data toe te lichten.

Samenwerking

Het onderzoek werd gestructureerd met de research data management lifecycle als leidraad. Voor elke fase van deze cyclus werd uitgevraagd hoe er met data wordt omgegaan en welke ondersteuning er aanwezig is in de organisatie om dit te faci-

teren. Dankzij steeds meer geavanceerde analytische technieken, komen er in de erfgoedwetenschappen ook steeds meer data beschikbaar. Goed data management is dan cruciaal. Echter is er momenteel nog geen uitgebalanceerde infrastructuur om dit te ondersteunen. E-RIHS (EU) wil met het DIGILAB zorgen voor een optimale data-infrastructuur en wil alle partners de mogelijkheid bieden om data, tools en expertise makkelijker uit te wisselen. Ook FAIR data is uiteraard belangrijk voor het netwerk. FAIR data maken een grotere impact en kunnen beter worden gedeeld en hergebruikt. Uit de landschapsstudie kwam duidelijk naar voren dat de partners enthousiast zijn over het E-RIHS.nl netwerk omdat dit netwerk een katalysator vormt voor

betere samenwerking.

Training

Naast de inventarisatie, meer gericht op de technische data-infrastructuur, verzorgde DANS ook een eerste training voor het E-RIHS.nl netwerk. Samen met het Netwerk Digitaal Erfgoed (NDE) werd deze zomer de bootcamp ‘Duurzame Data voor de Heritage Sciences’ gelanceerd om (aankomende) specialisten en data stewards te trainen op het gebied van research data management, specifiek voor de erfgoedwetenschappen. De unieke bootcamp was samengesteld uit modules en informatiebronnen uit de al bestaande trainingen Essentials 4 Data Support van RDNL en Leren Preserveren van het NDE. De deelnemers aan de bootcamp had-

den heel verschillende achtergronden en leerden over de verschillende inzichten, standaarden en terminologieën voor zowel erfgoeddata als onderzoeksdata. Belangrijk resultaat van de bootcamp was dat de cursisten ook vooral heel veel van elkaar leerden en elkaar begrijpen en elkaars data-taal leren spreken is een belangrijke focus voor verdere E-RIHS.nl trainingen.

Rapport

Alle bevindingen uit de landschapsstudie zijn gebundeld in een rapport dat aan de E-RIHS.nl werkgroep gepresenteerd is om de verdere oprichting van het netwerk te bevorderen. Het rapport lees je op: <https://e-rihs.nl/samenwerken-met-en-aan-data/>

Van het Rijksmuseum tot de Getty Provenance Index

Nederlandse datastandaard voor erfgoedcollecties

Iconclass is een belangrijke standaard voor de ontsluiting van beeldcollecties. De nieuwe online versie combineert zoeken op onderwerp met patroonherkenning. Classificeren kan nu sneller en beter en de rijke illustratie maakt het werken met het systeem ook een stuk aantrekkelijker. *Hans Brandhorst*

Ook als je bij een bezoek aan de site van het Rijksmuseum kiest voor de Nederlandse versie, blijft de informatie over het onderwerp van een schilderij hardnekkig een lijstje Engelse termen tonen. Zo zie je bij een topstuk als Frans Hals’ huwelijksportret van Isaac Massa en Beatrix van der Laen onder meer ‘wedding ring’ staan. De reden daarvoor is simpel: die Engelse begrippen zijn ontleend aan Iconclass en wie de cursor over het lijstje beweegt ziet dat inderdaad bevestigd in de statusbalk van de webbrowser.

Ontstaan in WOII

Het Iconclass-systeem is een metadata-standaard voor de inhoud van beelden, bedacht door Henri van de Waal (†1972), hoogleraar

kunstgeschiedenis in Leiden. Het basisidee van een alfanumerieke codering van iconografische begrippen ontstond al in de oorlogsjaren. Van de Waal voorzag bovendien al in de jaren ’50 dat geautomatiseerde dataverwerking een rol zou gaan spelen bij de toepassing van Iconclass.

Patroonherkenning

De meest recente toevoeging aan de nieuwe Iconclass Illustrated Edition maakt het mogelijk om, net als bij Google Images, een zoekactie naar een iconografisch thema te beginnen met een door de gebruiker zelf geüploade afbeelding. Die patroonherkenning van digitale beelden zou zelfs een visionair als Van de Waal zich niet hebben kunnen voorstellen.

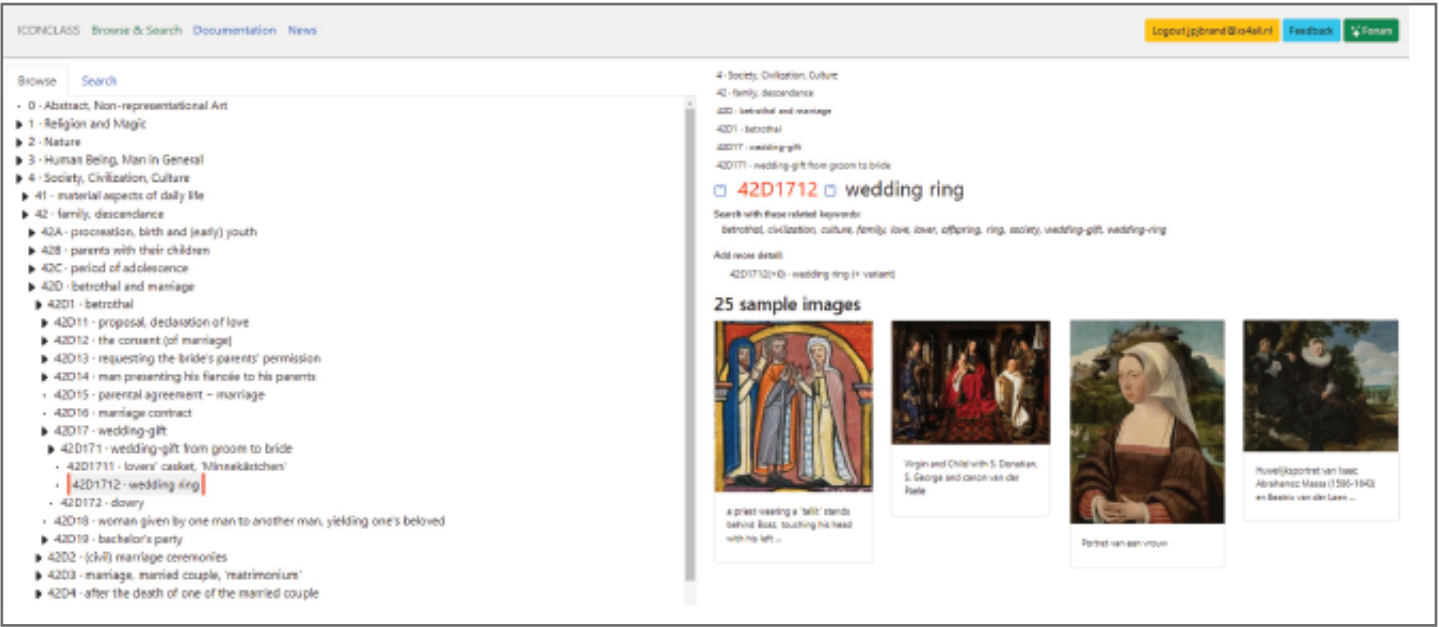
Dat je met de simpele code 42D1712 een ‘wedding ring’ in uiteenlopende erfgoedcollecties kunt terugvinden, zou hij volkomen logisch gevonden hebben. Die gestandaardiseerde codering zorgt er nu eenmaal vanzelf voor dat het niet uitmaakt of je het hebt over een Hochzeitsring, een alliance, een anello nuziale, een anel de casamento, of een 結婚指輪.

In eigen taal

Hij zou trots geweest zijn op de wereldwijde verspreiding van het systeem, maar zich wel hebben afgevraagd of musea niet meer zouden kunnen doen met die handige alfanumerieke codering. Daarmee kun je namelijk niet alleen trouwringen opsporen, maar ook alle afbeeldingen die met verloving en huwelijk in het algemeen te maken hebben, in één keer ophalen; en bovendien Japanse onderzoekers in hun eigen taal de metadata van musea laten doorzoeken... Maar alles overziend zou hij waarschijnlijk toch tevreden hebben vastgesteld wat hij, volgens overlevering, nuchter concludeerde over zijn eigen systeem: “Het werkt”.

De Henri van de Waal Stichting stelt de Iconclass Illustrated Edition beschikbaar aan iedereen die een beeldcollectie op een gestandaardiseerde wijze op onderwerp doorzoekbaar wil maken.

<https://iconclass.org/>
<https://henrivandewaalfoundation.org/>
Voor meer informatie:
info@iconclass.org



Weergave van zoekresultaat ‘wedding ring’ op de website van Iconclass

Unieke geluidsopnamen te beluisteren

Interviews van musea ontsloten in CLARIAH

Bij musea liggen vele waardevolle interviews over collectiestukken op de plank. Het project OH-SMART werkt aan een manier om deze te delen met de sector en de wetenschap. *Erica Renckens*

Interviews met kunstenaars over de gedachte achter hun oeuvre, met restaurateurs over de gemaakte afwegingen bij het herstellen van collectiestukken, met kunsthistorici over de betekenis van kunstwerken in een brede maatschappelijke context. Museale instellingen beschikken vaak over unieke geluidsopnames, waardevol en interessant voor collega’s uit de eigen sector en uit de wetenschap, maar niet vindbaar of toegankelijk. Het door CLARIAH gefinancierde project OH-SMART (Oral History - Stories

at the Museum around Artworks) moet hier verandering in brengen.

Benodigde standaarden

“Om audiomateriaal te kunnen hergebruiken, moet het op een goede manier zijn gearchiveerd en ontsloten. Daarvoor moeten nogal wat stappen worden gezet”, vertelt Femmy Admiraal, Chief Data Officer van CLARIAH en Data

Station Manager Humanities bij DANS. “De benodigde standaarden en procedures bestaan eigenlijk al jaren en beginnen nu ook hun plek te vinden in de Oral History.”

Virtuele werkomgeving

Binnen OH-SMART ontwikkelen onderzoekers van de Universiteit van Amsterdam, de Universiteit Twente en DANS een technische workflow

Stories in Motion

Admiraal is vanuit DANS ook betrokken bij het project Stories in Motion, waarin onder andere interviews met vrouwen die vanaf de jaren vijftig actief waren in de Rotterdamse vrouwenbeweging worden gearchiveerd. “Hier werken we niet aan een technische, maar aan een procedurele workflow. Dat is dus een beschrijving van de verschillende stappen die medewerkers van in dit geval het Stadsarchief moeten zetten om de interviewdata goed te archiveren. Dat gebeurt dan via de standaard interface hiervoor van DANS.” De data worden zo toegankelijk voor archivariissen, onderzoekers (via de CLARIAH Media Suite), journalisten en het grote publiek.



Onderzoekersrestaurator Sanneke Stigter interviewt tentoonstellingscurator Bas Mühren in het Kröller-Müller Museum

Credits: Marjon Gemmeke

waarmee musea hun opnames stap voor stap kunnen archiveren. Admiraal: “De data moeten in het juiste formaat staan, voorzien van een transcript dat is gegenereerd door de automatische spraakherkenner, zodat de opname doorzoekbaar wordt. Zo’n transcript is niet gelijk foutloos, dus dat moet nog bewerkt en gecorrigeerd worden. Al die stapjes zetten we nu op in een virtuele werkomgeving.” Dit voorjaar zijn de ontwikkelaars

begonnen aan het samenstellen van de virtuele werkomgeving. Een eerste versie is nu beschikbaar en die zal gedurende het OH-SMART project verder worden bijgeschaafd. “Idealiter wordt deze dan ook voorzien van een interface, zodat medewerkers van een museale instelling makkelijk alle stappen kunnen doorlopen en hun interviews kunnen archiveren bij DANS.”

<https://www.uva.nl/ohsmart>

Innovatieve projecten in het Centerpanel en LISS panel

Vragenlijstenonderzoek in de toekomst

Bijna alle onderzoeken die plaatsvinden in het LISS panel maken gebruik van vragenlijst-onderzoek: panelleden krijgen een vragenlijst voorgelegd over een actueel onderwerp en vullen deze in. Het LISS panel wordt echter steeds vaker gevonden door onderzoekers die via een niet traditionele manier data willen verzamelen onder de Nederlandse bevolking. Mara Verheijen



Credits: Pixabay

Op dit moment worden er meerdere pilotprojecten in het Centerpanel en LISS panel afgenomen, waaronder een studie waar panelleden gevraagd worden een export te maken van hun Google Locatie Data en een studie waar panelleden een coöperatief online spel spelen, tegelijk met andere panelleden. Deze studies laten zien dat in het LISS panel meer mogelijk is dan alleen vragenlijstenonderzoek.

Pilotstudie 1: Donatie van Google locatiedata

Veel mensen laten (bewust of onbewust) veel digitale sporen (persoonsgegevens) achter door het gebruik van hun mobiele apparaten. Deze digitale broodkruimels bieden interessante mogelijkheden voor sociaalwetenschappelijk onderzoek. Vanaf mei 2018 is elke entiteit die persoonsgegevens van inwoners van de Europese Unie verwerkt, wettelijk verplicht om deze persoonsgegevens op verzoek

aan gebruikers te verstrekken. Grotere organisaties zoals Google of WhatsApp bieden hun gebruikers daarom de mogelijkheid aan om hun gebruikersgegevens te downloaden via zogenaamde 'Data Download Pakketten' (DDP's).

Eenvoudig data doneren

In een studie met ruim 1100 deelnemende Centerpanel-leden, werd gevraagd of panelleden hun DDP met Google Location History data wilde doneren. Hiervan bleek ruim een kwart van de panelleden bereid te zijn om zijn of haar datapakket te doneren. Na instemming werden panelleden stap voor stap meegenomen in het datadonatie proces. Door deze uitgebreide instructies lukte het om zelfs de oudste deelnemers in het panel te bereiken: de oudste donateur was ruim 87 jaar oud.

Binnenkort onderzoek met WhatsApp

Deze pilotstudie laat zien dat het mogelijk is om digitale traceergegevens van panelleden (o.a. Google Locatiegegevens, WhatsApp-gegevens) op een verantwoorde manier te verzamelen voor onderzoek. Na de implementatie van deze software in het Centerpanel zal op korte termijn ook een onderzoek met WhatsApp data plaatsvinden in het LISS panel. Met deze methode kan o.a. de betrouwbaarheid van de digitale traceergegevens onderzocht worden als aanvulling op zelfrapportage. Ook kan er onderscheid gemaakt worden tussen deelnemers die bereid zijn om die digitale sporen te delen en degenen die dat niet zijn. Al met al biedt deze aanpak veel nieuwe mogelijkheden voor toekomstig onderzoek in het LISS panel.

Pilotstudie 2: Coöperatief onderhandelingspel

Een andere nieuwe software die sinds kort geïntegreerd is het Centerpanel en het LISS panel is de software van het platform oTree. oTree is een open-source op Python gebaseerd softwareplatform dat wordt gebruikt voor online interactief gedragsonderzoek en experimenten.

Onderhandelingspel geeft inzicht in gedrag

De pilotstudie werd gedaan door onderzoekers van Centerdata en Tilburg University voor een onderzoeksproject over menselijk gedrag bij coalitievorming. Voor dit project is een spel ontworpen waarbij LISS panelleden elk een transportbedrijf vertegenwoordigen. Nadat oTree automatisch meerdere groepjes van drie respondenten heeft gevormd, werken zij samen om hun lading van A naar B te vervoeren en te onderhandelen over de (feitelijk uitgedaande) winstverdeling.

Voor de pilotstudie werden 744 panelleden gevraagd om deel te nemen aan het onderhandelingspel. Een derde van de respondenten stemde in met deelname aan het experiment. Ongeveer 75% startte daadwerkelijk met het experiment, nog eens driekwart hiervan voltooide het spel. Het percentage spelers dat kon worden gegroepeerd en een coalitie kon vormen was 60%. De rest viel uit of strandde omdat één van de andere spelers afhaakte tijdens de onderhandelingen. De pilotstudie is succesvol afgerond: de lessen getrokken uit de pilot zijn gebruikt om de hoofdstudie in het LISS panel af te nemen. Op basis van aankomende studies wordt de integratie van realtime speltheorie-experimenten in het LISS panel verder geoptimaliseerd.

COLUMN

Open data: voor een transparante overheid!

Ik had 10 jaar geleden niet kunnen voorspellen dat ik nu bij het Nationaal Archief zou werken. Het imago van een archief als stoffige plek waar oude collecties worden ondergebracht domineert helaas nog steeds. Echter, het gezonde transparantie-activisme dat ik de voorbije jaren had ontwikkeld, was precies wat me naar het Nationaal Archief dreef. Ik geloof namelijk dat de kennis en expertise waar het Archief over beschikt, behulpzaam kan zijn om overheden transparanter te maken. Deze expertise wordt nu echter nog vaak ingezet aan de achterkant van de informatieketen: om collecties voor de eeuwigheid te bewaren en ontsluiten. Dat geldt ook voor (linked) open data. Inmiddels zijn zo'n 400.000 foto's uit onze collectie beschikbaar als open data. Ook zijn we

gestart om archieven over de Tweede Wereldoorlog als linked open data beschikbaar te stellen. Samen met anderen zoals het Netwerk Oorlogsbronnen, het Huygens instituut en het NIOD. Onlangs sprak ik iemand over een initiatief van de Open State Foundation, ondersteund door het Stimuleringsfonds voor de Journalistiek. De Open State Foundation heeft een online platform ontwikkeld waarin lokale overheidsdata, zoals raadsbesluiten en gemeentelijke uitgaven, zijn gekoppeld en gebruiksvriendelijk doorzoekbaar gemaakt voor journalisten of bezorgde burgers.

De maatschappelijke roep om een open en transparantere overheid klinkt steeds luider. Denk aan de kindertoeslagenaffaire, de gaswinning in Groningen

en de vele WOO-verzoeken. Tegelijkertijd is het voor veel overheden een uitdaging om aan deze openheid te voldoen. Het is vaak zoeken naar een speld in een digitale hooiberg. Precies daarom gaat het Nationaal Archief zijn expertise intensiever inzetten voor rijksoverheden, aan de voorkant van de informatieketen.

Hoe mooi zou het zijn wanneer we het initiatief van de Open State Foundation op landelijk niveau een vervolg kunnen geven? Dat we met elkaar een magneet ontwikkelen waarmee iedereen binnen no time de voor hen relevante informatie uit die ontzaglijke informatieberg van de overheid kan trekken. En niet alleen uit de collecties van het Nationaal Archief, maar ook de vele andere archieven in het land. Om

dat te realiseren moeten we beginnen bij de inrichting van processen en systemen bij die overheden, aan de voorkant van de informatieketen. Om daar de toepassing van linked open data in te regelen.

Heb je andere ideeën of suggesties hoe we met open data én elkaar die magneet kunnen bouwen? Neem dan vooral contact met ons op via contact@nationaalarchief.nl. Mijn medewerking heb je!



Zo dragen we gezamenlijk bij aan een open en transparante overheid.

Tom de Smet

Tom de Smet is sinds oktober 2021 directeur Archieven, Dienstverlening en Innovatie, en tevens CIO, bij het Nationaal Archief. Na zijn studies Psychologie in het Verenigd Koninkrijk, bracht zijn liefde voor film en cultuur hem naar Venetië waar hij filmfestivals produceerde. Van 2019 tot 2021 transformeerde hij als algemeen directeur het voormalige Communicatiemuseum aan de Zeestraat in Den Haag (nu: Beeld en Geluid Den Haag) tot een museum waar nieuwswijsheid, 'fake news', en het belang van een vrije pers in de informatievoorziening aan burgers centraal staat. Hij geeft de column door aan Gerard van der Steenhoven.