

Onderzoek voor de Wereldbank

Hoe beweegmeters inzicht bieden in het leven in Malawi

Hoe onderzoek je dagelijkse tijdsbesteding zonder mensen elke dag vragenlijsten te laten invullen? Data scientists van Centerdata ontwikkelden een manier om met machine learning dagelijkse bezigheden te herkennen uit data van beweegmeters. In opdracht van de Wereldbank testen ze daarvoor nu een nieuw, verbeterd model met data verzameld op het platteland van Malawi. *Mara Verheijen*



Leven op het platteland. Bron: Adobestock

Innovatief onderzoek met beweegmeters

Onderzoek naar tijdsbesteding is niet nieuw. Meestal worden hiervoor vragenlijsten gebruikt. Respondenten moeten dan bijvoorbeeld aan het eind van de dag opschrijven wat ze die dag gedaan hebben. Aan die onderzoeksmethode kleven echter nadelen: het kost veel tijd en levert niet altijd de beste data op. Dat geldt zeker bij respondenten die moeite hebben met lezen en schrijven of waarvoor het dagritme niet zo duidelijk bepaald wordt door de tijd op de klok. Ook onderzoek met beweegmeters is niet nieuw. Onderzoekers geïnteresseerd in gezondheid proberen regelmatig met draagbare sensoren of smartwatches te achterhalen of mensen voldoende beweging krijgen. Wat wél nieuw is, is de combinatie: tijdsbestedingsonderzoek met beweegmeters. Die combinatie lijkt een veelbelovende manier te bieden om meer inzicht te krijgen in tijdsbesteding in delen van de wereld waar weinig mensen een opleiding gevolgd hebben en de inkomens laag zijn.

Kunnen beweegmeters vragenlijsten vervangen?

Inzichten uit dergelijk onderzoek zijn van be-

lang voor beleidsmakers. De Wereldbank vroeg Centerdata daarom al eerder om te kijken naar de mogelijkheden om op een eenvoudige, objectieve en goedkope manier inzichten te verkrijgen met beweegmeters. In 2019 startte een groot onderzoek in Malawi. In dat onderzoek vulden zo'n 400 mensen uit agrarische gemeenschappen in Malawi vragenlijsten in. Daarnaast werden hun activiteiten gemeten met beweegmeters. En hoewel de data uit beweegmeters niet zomaar af te lezen zijn, lukte het om een machine learning model te ontwikkelen dat op basis van de data voorspelde met welke activiteit iemand bezig was. In dat onderzoek kwamen al duidelijke verschillen tussen mannen en vrouwen naar voren. Onderzoek met beweegmeters bleek dus een kansrijke manier om meer te weten te komen over tijdsbesteding.

Verdere ontwikkeling van het voorspellingsmodel

In het huidige project zetten we in op nog betere inschattingen van wat mensen doen. De Wereldbank verzamelde daarvoor opnieuw data in Malawi. Nu werden bij ruim 1440 mensen onder andere hoogfrequente sensordata verzameld. Dat wil zeggen dat in plaats van één meting per minuut er dit keer 60 metingen per seconde werden gedaan. Daarmee verbeteren onze data scientists het voorspellingsmodel. Zo hopen we uiteindelijk nog meer en preciezer inzicht te krijgen in de leefomstandigheden van mensen in ontwikkelingslanden én de techniek van tijdsbestedingsonderzoek met beweegmeters verder te ontwikkelen.

centerdata.nl

CBS werkt aan internationaal data delen

Het Centraal Bureau voor de Statistiek (CBS) is lid van het International Data Access Network (IDAN), een samenwerkingsverband tussen vijf Europese onderzoeksinstituten gericht op het faciliteren van statistisch-wetenschappelijk onderzoek. Deze onderzoeksinstituten werken aan het vergemakkelijken van toegang tot hun data vanuit het buitenland. *Maarten Bloem*

Het CBS draagt bij aan deze inspanning door microdata beschikbaar te stellen voor statistisch-wetenschappelijk onderzoek aan onderzoekers uit de Europese Economische Ruimte en landen met een adequaatheidsbesluit van de Europese Commissie (EC). Dit betekent dat deze landen volgens de EC een passend niveau van bescherming van persoonsgegevens bieden, ofwel door nationale wetgeving of op basis van internationale

verplichtingen. Deze faciliteiten worden aangeboden als een 'remote access'-voorziening, wat betekent dat onderzoekers vanuit het buitenland, eenmaal geautoriseerd, toegang kunnen krijgen tot een beveiligde microdata-omgeving op hun eigen computer. Dit in tegenstelling tot veel andere Europese statistiekbureaus en datahouders die specifieke eisen stellen aan apparatuur of een daarvoor bestemde ruimte, wat fysieke reizen vereist voor onderzoekers die data uit andere landen willen bestuderen. De leden van IDAN, waaronder naast het CBS het Franse Centre d'accès sécurisé aux données (CASD), het Duitse Institut für Arbeitsmarkt- und Berufsforschung (IAB), het Leibniz Institute for the Social Sciences (GESIS) en het Britse UK Data service (UKDS), hebben in 2017 gezamenlijk standaarden ontwikkeld voor toegang tot gepseudonimiseerde microdata. Dit resulteerde in de

oprichting van 'access points' waar onderzoekers toegang kunnen krijgen tot data in een beveiligde ruimte die voldoet aan technische en juridische eisen. Momenteel heeft het CBS een access point gerealiseerd voor data van CASD, en er wordt gewerkt aan de totstandkoming van access points voor data van UKDS, IAB en GESIS. Dit opent nieuwe mogelijkheden voor het CBS en onderzoekers, die toegang krijgen tot een diverse reeks buitenlandse data, waardoor de kansen voor statistisch onderzoek aanzienlijk toenemen. Elke instelling binnen IDAN hanteert specifieke procedures voor het aanvragen van toegang tot data. Op de website van IDAN is een overzicht van de access points te vinden, evenals links naar relevante informatie over beschikbare data en aanvraagprocedures van alle deelnemende instellingen. idan.network/#section-1



E-DATA & RESEARCH

Jaargang 18 | nummer 2

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

E-data & Research verschijnt drie keer per jaar en wordt mogelijk gemaakt door: CBS, Centerdata, CLARIAH, DANS, KNAW Humanities Cluster, ODISSEI en RKD.

INHOUD

3 RKD Research maakt kunsthistorisch onderzoek makkelijker

3 Waardevolle data delen

4 Databestanden sinds kort beschikbaar bij CBS, Centerdata en DANS

6 SSHOC-NL is de nieuwe onderzoeksinfrastructuur

7



HisGIS plakt historische kaarten op het moderne landschap

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

GEHOORD & BIJGEWOOND

Symposium ‘Oral History, een stap verder’

Widia Mahabier

Op vrijdag 8 december jl. vond het symposium ‘Oral History, een stap verder’ plaats, bij de Rijksdienst voor het Cultureel Erfgoed (RCE) in Amersfoort, georganiseerd door het Knooppunt ‘Sprekende geschiedenis’.

Het was een mooie en waardevolle bijeenkomst waarin oral history in al zijn gelaagdheid is besproken en doordacht, en er is veel uitgewisseld. Dagvoorzitter en initiatiefnemer van ‘Sprekende geschiedenis’ Saskia Moerbeek trapte de dag af waarna de eerste spreker het podium betrad. James Kennedy, hoogleraar moderne Nederlandse geschiedenis aan de Faculteit Geesteswetenschappen van de Universiteit Utrecht, sprak over de diepgaande impact van oral history op individuen en gemeenschappen. Een tweede presentatie werd gegeven door Rose Mary Allen, bijzonder hoogleraar Cultuur, gemeenschap en geschiedenis aan de Universiteit van Curaçao. Zij sprak onder andere over haar ervaringen met oral history op de Cariben. Frank von Meijenfeldt van het Knooppunt ‘Sprekende geschiedenis’ vertelde over de laatste technologische ontwikkelingen op het gebied van oral history. Daarna konden de deelnemers verschillende workshops volgen.

Met een interactieve demonstratie toonden DANS-ers Jetze Toubert en Widia Mahabier in de workshop ‘Bewaren en toegankelijk maken van oral history interviews’ de

mogelijkheden voor het duurzaam bewaren en toegankelijk maken van oral history interviews in het DANS Data Station Social Sciences and Humanities. Tijdens de afsluitende bijeenkomst zijn de resultaten van de verschillende workshops plenair gedeeld.

https://sprekendeschiedenis.nl/symposium_2023/

CLARIAH Annual Conference, Paushuize, Utrecht

Jetze Toubert

30 november jl. hield CLARIAH, de nationale infrastructuur voor digitale onderzoek in de geesteswetenschappen, zijn jaarlijkse conferentie. Deze editie was de afsluiting van de huidige fase van CLARIAH, dat tot nu toe op projectbasis opereert. Het thema van de conferentie was dan ook ‘past, present and how we shape our digital humanities future’: zowel terugkijken op de behaalde resultaten, als stilstaan bij de huidige stand van zaken, én vooruitkijken naar toekomstige ontwikkelingen. De ochtend stond in het teken van weidse perspectieven en vergezichten. Aan bod kwamen de plaats die CLARIAH in dertien jaar tijd heeft verworven in het landschap van onderzoeksinfrastructuur (Dirk van Miert, Principal Investigator CLARIAH), de technische voorzieningen die in CLARIAH zijn ontwikkeld (Roeland Ordelman, Chief Technology Officer CLARIAH), de integratie van digitale faciliteiten in het geesteswetenschappelijke onderwijs (Susan Aasman, Rijksuniversiteit Groningen), en de digitale stand van zaken in het on-

derzoek in de humaniora (Joris van Eijnatten, eScience Center). De ochtend werd afgesloten met een paneldiscussie, waarin Julia Noordergraaf (Universiteit van Amsterdam) nog aansloot. In de discussie kwam naar boven dat voortdurend kritisch moet worden bekeken waar het priemaat ligt, technologische ontwikkelingen of de behoeftes van onderzoekers. Ook werd benadrukt dat digitale vaardigheden net zo belangrijk zijn om te cultiveren als geavanceerde infrastructuur, en dat geesteswetenschappers vooral moeten leren in teamverband data intensief onderzoek op te pakken. Deze thema’s kwamen in het middagprogramma terug, waarin in vijf parallelsessies interactieve workshops werden gehouden, over onderwerpen als de Media Suite (een geïntegreerde onderzoeksomgeving met data en tools voor geesteswetenschappers), het vergroten van de herbruikbaarheid van data in de geesteswetenschappen (onder meer via de DANS Data Stations), de faciliteiten om research software te ontwikkelen in samenwerking met het eScience Center, de mogelijkheden van knowledge graphs om (zeer) grote hoeveelheden data te structureren en te verwerken, en de toepassing van extended reality en artificial intelligence in de geesteswetenschappen.

Het was een geslaagde dag met veel geanimeerde gesprekken, waar zowel oudgedienden in het CLARIAH ecosysteem als nieuw geïnteresseerden volop aan deelnamen.

<https://www.clariah.nl/events/clariah-annual-conference-2023>

ODISSEI Conference for Social Science in the Netherlands 2023

Ricarda Braukmann

Op 2 november 2023 werd in Utrecht de jaarlijkse ODISSEI Conferentie gehouden waarin de ontwikkelingen van digitale en computationele methodes in sociaal-wetenschappelijk onderzoek centraal stonden.

Linnet Taylor, Professor voor International Data Governance bij het Tilburg Institute for Law, Technology, and Society, startte de dag met een kritische reflectie op het toenemende gebruik van grote en gelinkte datasets binnen de sociale wetenschappen. Aan de hand van verschillende voorbeelden werden de ethische vragen en praktische overwegingen duidelijk die de onderzoeksgemeenschap zich moet stellen in het gebruik van big data en nieuwe technologieën.

Na de keynote werd het tempo verhoogd door middel van een sessie flash talks. In dit experimentele format kregen vijf sprekers de mogelijkheid om in vijf minuten hun verhaal te doen. Zo werden SANE, de ODISSEI Portal, de DANS Data Station SSH, het SoDA Team en een showcase van het gebruik van CBS microdata in combinatie met panel data in korte tijd uitgelicht.

Wie meer wilde weten over deze faciliteiten kon in de pauze langslopen bij de stands van de ODISSEI Marketplace. Hier waren ook het LISS panel en het Thematic Digital Competence Centre SSH vertegenwoordigd. Bezoekers konden ook de 36 poster bekijken die door veelal jonge onderzoekers werden gepresenteerd.

Daarna konden de ruim 300 deelnemers zich verdelen over drie parallelsessies met elk vijf onderwerpen. Wie zich wilde verdiepen in methodes kon sessies bijwonen over smart data collection, agent based modelling, machine learning en benchmarking. Thematische sessies werden gehouden rondom misinformatie en gezondheidsdata en er waren sessies over population network, tekstanalyse en een aantal sessies gericht op innovatie in onderzoeksinfrastructuur.

In de tweede keynote vertelde David Lazer, University Distinguished Professor Political Science and Computer Sciences bij de Northeastern University, over de recente ontwikkelingen in het gebruik van sociale media data - met name Twitter - die steeds ontoegankelijker wordt voor onderzoek. Ook deze keynote bood een kritische blik op het onderzoeksveld en Lazer deelde veel van zijn ervaringen en inzichten over paradigma’s voor het bestuderen van online gedrag van mensen en platforms.

Na de keynote werd de posterprijs uitgereikt aan Lydia Repke, postdoc bij GESIS Leibniz Institute for Social Sciences. Voor de welverdiende borrel werd het concept van de flash talks nog een keer herhaald. Dit keer mochten vijf jonge onderzoekers hun studies pitchten. Een hele leuke manier om kort te horen over een project en vervolgens het gesprek op te zoeken bij de borrel om het verder te kunnen verdiepen.

<https://odissei-data.nl/>

NPSO innovatie dag

Mara Verheijen

Op 29 november 2023 organiseerde de NPSO (Nederlandstalig platform voor Survey Onderzoek) haar tweejaarlijkse Innovatiedag. Deze 6e

editie vond plaats bij het CBS in Den Haag. Tijdens de Innovatie dag presenteerden 14 PhD studenten uit Nederland en België hun onderzoek. De onderzoeksonderwerpen die aan bod kwamen waren divers. Zo werd er gepresenteerd over non-probability samples, synthetische data, het laten doneren van verzamelde gegevensdata, verschillende push-to-smart-strategieën en nog veel meer. Een slotlezing werd gegeven door Joost Kappelhof, hoofd methodologie van het Sociaal Cultureel Planbureau.

PhD studente Inge Wolf (Amsterdam UMC) won met haar presentatie over haar onderzoek de NPSO innovatieprijs. In haar onderzoek maakt Wolf gebruik van beweegmeters om fysieke activiteit te meten. Dit deed ze onder andere in het LISS panel van Centerdata.

Meer weten over NPSO?
Het NPSO wil iedereen in Nederland en Vlaanderen samenbrengen die zich bezighoudt met de methodologie van het uitvoeren van grootschalig sociaal-wetenschappelijk onderzoek onder de bevolking. Onderzoekers kunnen afkomstig zijn uit overheden, universiteiten en andere onderzoeksinstituten en het bedrijfsleven. Vanaf 2008 organiseert het NPSO een vaste, jaarlijkse lezingendag over specifieke thema’s en meerdere keren per jaar een middag over een specifiek onderwerp. Al deze bijeenkomsten zijn gratis voor deelnemers, en worden gesponsord door de instituten die vertegenwoordigd zijn in het NPSO. Voor aankomende evenementen, zie de website van het NPSO:

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

AGENDA

27 februari • Utrecht

Teaming up across domains

Als we echt vooruitgang willen boeken op het gebied van FAIR data, open science en onderzoekssoftware, moeten we samenwerken en de wetenschappelijke en professionele domeinen overstijgen.

tdcc.nl/evenementen/learning-up-across-domains/

29 februari • Utrecht

Essentials 4 Data Support

Een introductiecursus voor diegenen die onderzoekers (willen) ondersteunen bij het opslaan, beheren, archiveren en delen van hun onderzoeksdata.

datasupport.researchdata.nl/

7 maart • Den Haag

Microdatamiddag CBS

‘Bestaanszekerheid’

Keynote door CBS hoofdeditor Peter Hein van Mulligen. Er is plek voor maximaal 80 deelnemers. Bij een hoger aantal inschrijvingen wordt gewerkt met een wachtlijst.

Let op! Voertaal is Nederlands. cbs.nl

18 - 21 juni • Lissabon

DARIAH Annual Event 2024

De conferentie richt zich op de technische, methodologische infrastructuurle en conceptuele uitdagingen van het ontwerpen, implementeren en delen van digitale ondersteunende workflows. dariah.eu/event/dariah-annual-event-2024/

De digitale transformatie bij het RKD – Nederlands Instituut voor Kunstgeschiedenis

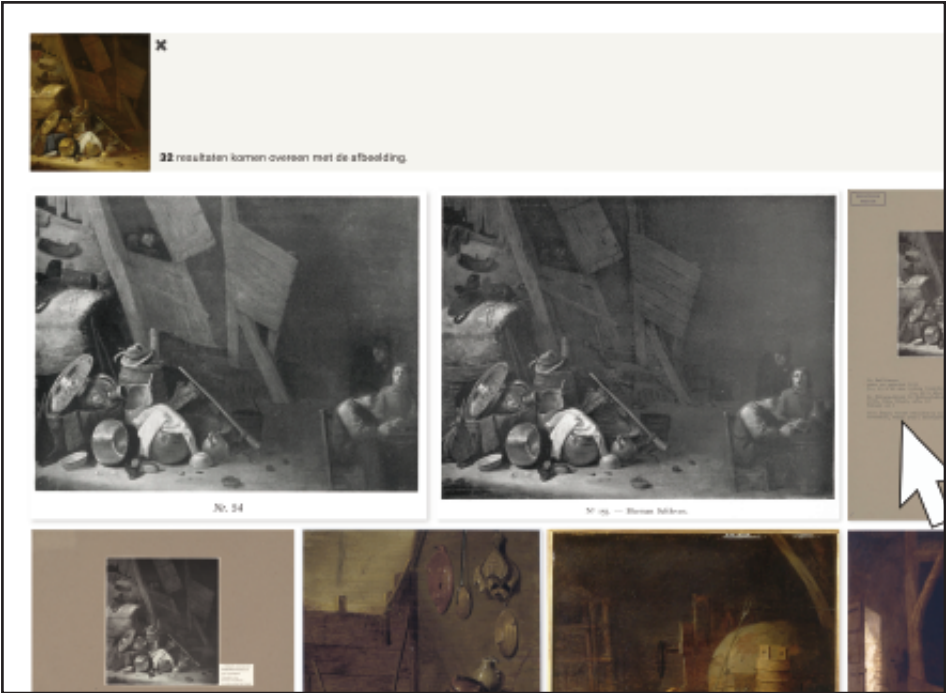
RKD Research maakt kunsthistorisch onderzoek makkelijker

Veel onderzoekers maakten al gebruik van de databases en het onderzoeksmateriaal op de studiezaal van het RKD – Nederlands Instituut voor Kunstgeschiedenis. Nu zijn er nog meer mogelijkheden voor digitaal onderzoek. *Marjon van Schendel*

Vanaf deze maand zijn miljoenen extra data online gekomen op het nieuwe platform RKD Research, dé digitale entree voor kunsthistorisch onderzoek. Daarbij biedt het innovatieve visual search ongekend veel mogelijkheden tot kunsthistorische ontdekkingen.

Digitale transformatie

Het RKD is een van de belangrijkste kunsthistorische kennisinstellingen ter wereld. Het instituut verzamelt, beheert, onderzoekt en presenteert een unieke rijkscollectie, waaronder meerdere databases, honderden archieven en 's werelds grootste bibliotheek met betrekking tot beeldende kunst van de Nederlanden in internationale context, van de late middeleeuwen tot heden. Kunsthistorici zijn gewend om gedurende hun onderzoek langs te komen op de studiezaal van het RKD. Binnen het RKD leefde al langer de gedachte dat door digitalisering veel meer onderzoekers, ook internationaal, in de gelegenheid gesteld kunnen worden om bij het onderzoeksmateriaal te komen, waardoor nieuwe mogelijkheden voor verbinding van de gegevens ontstaan. Inmiddels is de volledige beelddocumentatie van het RKD gedigitaliseerd, daardoor zijn



Met visual search kun je in de databases zoeken op afbeelding.

ruim vier miljoen afbeeldingen van kunstwerken met onderzoeksgegevens digitaal beschikbaar gekomen. Hier eindigde de bouw van RKD Research niet. Voor digitaal onderzoek is het ook van belang dat de metadata op een andere manier beschikbaar gesteld worden. In lijn met de Nationale Strategie Digitaal Erfgoed heeft het RKD al zijn databases, waaronder RKDimages, RKDartists en RKDtechnical als Linked Open Data (LOD) beschikbaar gesteld. Hierbij is onderzocht welke LOD ontologieën het best passen bij de gegevensverzamelingen. Voor de basis, de kunstwerkgegevens, is het Linked Art model gebruikt, voor RKDlibrary is RDA gebruikt en voor RKDarchives RICO. RKDartists is op basis van CIDOC-crm beschikbaar gesteld en de kern van de data wordt daarnaast nog steeds vol-

gens de schema ontologie aangeboden. Hiermee zijn ook de metadata van het RKD breder beschikbaar als Linked Open Data. Door het open karakter van de data ontstaat de mogelijkheid om verbindingen aan te gaan met andere type datacollecties en kunnen nieuwe onderzoeksvragen beantwoord worden.

Visual search

Door de digitale beschikbaarheid van de beelddocumentatie ontstaat de mogelijkheid om op een andere manier onderzoek te doen. Hiervoor heeft het RKD op basis van een tool van de universiteit van Oxford een mogelijkheid voor visueel zoeken ontwikkeld. Je kunt hiermee op afbeelding zoeken in de databases van RKD Research. Deze tool, visual search, biedt verrassende resultaten en

legt verbindingen die eerder door het fysiek doorzoeken van de mappen met beeldmateriaal niet te vinden waren. Als er op een kunstwerk wordt gezocht, komen bijvoorbeeld ook historische foto's en kopieën tevoorschijn. De tool zoekt ongeacht kleur en vindt ook fragmenten en gespiegelde afbeeldingen. Daarnaast is er de mogelijkheid om de afbeeldingen op te slaan en te bewerken in een eigen omgeving: my workspace. Deze biedt allerlei handige mogelijkheden om afbeeldingen met elkaar te vergelijken, in te zoomen en contrast en kleur aan te passen. Zo is bijvoorbeeld een signatuur op een schilderij beter te bekijken. Ook is het mogelijk om afbeeldingen zonder copyright op hoge resolutie te downloaden in plaats van dit aan te vragen bij het RKD.

Toegankelijkheid

RKD Research is voor iedereen vrij toegankelijk. De genoemde tools en het downloaden van afbeeldingen op hoge resolutie zijn alleen beschikbaar als je een abonnement neemt. Bekijk het platform op:

<https://research.rkd.nl>



Het onderzoeksplatform RKD Research is voor iedereen vrij toegankelijk.

Ervaringen van de YOUTH cohort studie

Complexe data toegankelijk maken en houden

De dataverzameling van de YOUTH cohort studie, een grootschalig Utrechts onderzoek dat kinderen volgt in hun ontwikkeling is inmiddels gestopt, maar de betrokken wetenschappers blijven zich inzetten om de waardevolle data toegankelijk te houden. *Ricarda Braukmann*

De YOUTH cohort studie is een omvangrijk, longitudinaal onderzoek waarbij de ontwikkeling van bijna 4.000 Nederlandse kinderen en hun ouders is gevolgd, vanaf de zwangerschap tot de vroege volwassenheid. Chantal Kemner, hoofdonderzoeker en hoogleraar Biologische ontwikkelingspsychologie in Utrecht, kijkt trots terug op het project dat tien jaar geleden van start ging: "YOUTH is echt uniek vanwege de fijnmazige manier waarop de ont-

wikkeling van sociale competentie en gedragscontrole in kinderen gevolgd wordt, met een grote rol voor neurocognitieve maten." De verzameling omvat een breed scala aan informatie, waaronder biomedische data, lichaamssamples zoals haar en bloed, MRI- en echo-beelden, antwoorden op vragenlijsten, video-opnamen van gedrag en interacties, en eyetracking-gegevens, die samen inzicht bieden in de fysieke en psychologische ontwikkeling van de kinderen.

Waardevolle data delen

Al vanaf het begin van het project werd volop ingezet op de veiligheid en de FAIR-heid (vindbaar, toegankelijk, interoperabel en herbruikbaar) van de data. De onderzoekers wilden ervoor zorgen dat de data breed gedeeld kunnen worden zodat zo veel mogelijk onderzoekers

ervan gebruik kunnen maken. YOUTH ontving voor deze inzet dan ook de Dutch Data Prize in 2022.

Veilige opslag

YOUTH werd gefinancierd door een Zwaartekrachtsubsidie van NWO voor het consortium on individual development (CID) dat inmiddels is afgerond. Daarmee is de dataverzameling ook afgerond en is de dataset compleet. "Het werkt stopt hier echter niet", vertelt Kemer. "Het FAIR en toegankelijk houden van data vraagt om blijvende inspanning van een team van onderzoekers en dat werk gaat door." Binnen het project heeft YOUTH veel geïnvesteerd in infrastructuur, waaronder het Yoda-systeem. Dit systeem biedt een veilige opslagoplossing voor de onderzoeksdata en is met medewerking van YOUTH ontwikkeld door de IT-

en bibliotheekdiensten van de Universiteit Utrecht. Daarnaast is er in latere fasen gewerkt aan de creatie van een metadataportal. Dit initiatief, tot stand gekomen in samenwerking met vijf andere grote cohortstudies, is ontworpen om aan te sluiten op de ODISSEI Portal-omgeving. Het uiteindelijke doel is om meer sociaal-wetenschappelijke data vindbaar en toegankelijk te maken voor onderzoekers.

Data hergebruik

Er is ook een aanvraagmodule ontwikkeld binnen Yoda, samen met YOUTH, zodat andere onderzoekers een aanvraag kunnen indienen om gebruik te maken van de data. Kemner: "Momenteel is er gelukkig veel aandacht voor data hergebruik maar het maken en bemensen van dit soort infrastructuur is heel kostbaar en het is zorgelijk dat de financiering hier-

voor vaak tijdelijk en versnipperd is." Om een dataset echt herbruikbaar te maken blijven experts nodig die het hergebruik kunnen begeleiden. Daar zou volgens Kemner nog meer aandacht voor moeten zijn: "Duurzaam investeren betekent niet alleen ervoor zorgen dat gebruik technisch mogelijk is, maar ook ervoor zorgen dat de dataset daadwerkelijk optimaal wordt gebruikt." FAIR en open data hebben dus ondersteunende wetenschappers nodig – die daarvoor daadwerkelijk de mogelijkheden krijgen zodat ook in de komende jaren nog veel onderzoekers gebruik kunnen blijven maken van de rijke en waardevolle data uit YOUTH. <https://www.uu.nl/onderzoek/dynamics-of-youth/onderzoek/het-youth-onderzoek>

SINDS KORT BESCHIKBAAR

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

Wind,
Windkracht,
Windenergie
Bron: Pixabay



CBS

• **Enquête met informatie over zelfstandigen in de zorg- en welzijnssector (AZWZZP).** Met de enquête wordt informatie verzameld over zelfstandig ondernemers zonder personeel die werkzaam zijn in de zorg- en welzijnssector. Het gaat daarbij om informatie over arbeidskenmerken zoals het aantal uren dat men werkzaam is en het beroep dat men heeft. Daarnaast gaat het over de redenen waarom ze werkzaam zijn als zelfstandige en de thema's scholing en vaardigheden en gezond en veilig werken.

<https://www.cbs.nl/nl-nl/onze-diensten>

• **Add-on geneesmiddelen en Overig Zorgproduct-stollingsfactoren verstrekt in de Medisch Specialistische Zorg (MSZDGAddonGeneesmVEKTTAB)** Dit bestand maakt onderdeel uit van een set van bestanden die betrekking hebben op behandelingen binnen de medisch specialistische somatische zorg (MSZ). Er is nu één bestand met add-on geneesmiddelen en OZP-stollingsfactoren (ook wel 'Dure Geneesmiddelen' genoemd) verstrekt in het verslagjaar (MSZDGAddonGeneesmVEKTJJJTABVV).

<https://www.cbs.nl/nl-nl/onze-diensten>



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

Centerdata

• **'Groener dan anderen?' Misvattingen over relatieve milieubetrokkenheid** In een eerdere versie van E-data (juni 2022) werd beschreven welke door ODISSEI gehonoreerde onderzoeksprojecten een data-verzameling in het LISS panel mochten afnemen. Een van die projecten ging over dat veel mensen zichzelf als 'groener' (meer betrokken bij het milieu) zien dan anderen. Dit project onderzoekt of mensen echt rapporteren en geloven dat ze groener zijn dan anderen en of deze (mis)percepties verschillen tussen groepen. Dit project onderzoekt ook of het corrigeren van mogelijke misvattingen over anderen milieuvriendelijke actie kan bevorderen.

lissdata.nl

DOI: <https://doi.org/10.17026/dans-zgq-ezva>

Ook sinds kort beschikbaar:

Studies LISS panel

- Barrett, J – September 2021 – Financial wellbeing. DOI: <https://doi.org/10.57990/vva6-b905>
- Bosma, F., van der Heide, A – June 2022 – Euthanasia and assisted suicide. DOI: <https://doi.org/10.57990/dqsk-x585>
- De Beurs, E – February 2022 - Life Events Checklist (LEC-5) & PTSS Checklist (PCL-5). DOI: <https://doi.org/10.57990/9vdm-2b07>
- De Beurs, E – February 2022 - The Inventory of Personality Organization (IPO-16-NL). DOI: <https://doi.org/10.57990/qeqv-v468>
- Koetse, M., Liu, Z., Bouma, J – Choice experiment in meat consumption. DOI: <https://doi.org/10.57990/t3wt-g135Mekkelholt, B., Roumen, J – July 2021 - Clarity of quotations. DOI: https://doi.org/10.57990/shxa-ba63>
- Mascini, P. Weber, F – July 2022 – Social Media experiences. DOI: <https://doi.org/10.57990/1b6a-8t92>
- Mastop, J., Bredenoort, D – December 2021 - About door-to-door sales. DOI: <https://doi.org/10.57990/qs2k-qt59>
- Ramaekers, M – February 2022 – Helping others. DOI: <https://doi.org/10.17026/dans-zpc-742n>
- Reep, C., Hupkens, C., Derksen, E – June 2022 – Eating habits. DOI: <https://doi.org/10.57990/y3g2-wb80>
- Reeskens, T., Metz, A – October 2021 - Personal Values and Covid-19. DOI: <https://doi.org/10.57990/zexf-8y56>
- Rekker, L – May 2022 - Heating Systems. DOI: <https://doi.org/10.57990/xs38-rr38>
- Tiemstra, T – February 2022 - Confidence in banks. DOI: <https://doi.org/10.57990/vyzd-rp22>
- Van der Torn, J., Holscher, M – July 2021 – Identity registration at work. DOI: <https://doi.org/10.57990/k67b-aq96>
- Vink, M., van der Toorn, J., Jansen W - September 2021 - Diversity in the Workspace. DOI: <https://doi.org/10.57990/ac31-t550>
- Vromans, R – April 2021 – Trauma Care. DOI: <https://doi.org/10.57990/dwm2-cr77>



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 in de DANS Data Stations:

- T. van der Meer; E. Steenvoorden; E. Ouattara (University of Amsterdam), 2023, 'Panel Survey - Covid Trust and Democracy (NL 2020)',

- <https://doi.org/10.17026/SS/WBN1Y5>, DANS Data Station Social Sciences and Humanities, V1
- I.A. Niet (Eindhoven University of Technology), 2023, 'EU public values in energy and IT policies', <https://doi.org/10.17026/SS/LXJPMT>, DANS Data Station Social Sciences and Humanities, V1
- J. Jansen (Leiden University), 2023, 'Executive functioning in antisocial behavior', <https://doi.org/10.17026/SS/DDNIZV>, DANS Data Station Social Sciences and Humanities, V2
- H. Scholten (University of Twente), 2023, 'Dataset of 'A diary study investigating the differential impacts of Instagram content on youths' body image', <https://doi.org/10.17026/SS/7M90LJ>, DANS Data Station Social Sciences and Humanities, V1
- M. Spruit; D. Oosting; C. Kreffer (The Hague University of Applied Sciences), 2023, 'Questionnaire for research into secure behaviour with mobile digital devices', <https://doi.org/10.17026/SS/PD54TO>, DANS Data Station Social Sciences and Humanities, V1
- L. Derksen (Universiteit Groningen), 2023, '3FM coding data', <https://doi.org/10.17026/SS/KJCTIE>, DANS Data Station Social Sciences and Humanities, V1
- A. Muja (ResearchNed), 2023, 'Startmonitor studiejaar 2022-2023', <https://doi.org/10.17026/SS/5YI3PW>, DANS Data Station Social Sciences and Humanities, V1
- M. Savelkoul; P. van Groenestijn; P. Scheepers; R. Loh; R. Eisinga, 2023 (Radboud University / RSCR), 'Social and Cultural Developments in the Netherlands 2022', <https://doi.org/10.17026/SS/IRHVOA>, DANS Data Station Social Sciences and Humanities, V1
- E. Jacobs (Oberon), 2023, 'Schoolkostenmonitor po vo mbo 2022-2023', <https://doi.org/10.17026/SS/9PWKH6>, DANS Data Station Social Sciences and Humanities, V2
- S. Kliphuis (Utrecht University), 2023, 'Effects of lighted incubation and foraging enrichment during rearing on individual fear behaviour, corticosterone, and neuroplasticity in laying hen pullets - Data and protocols behaviour tests', <https://doi.org/10.17026/LS/OEMOV>, DANS Data Station Life Sciences, V2
- L. van Buul (Amsterdam UMC), 2023, 'Antibiotic prescribing and Non-prescribing in Nursing home residents with signs and symptoms Ascribed to urinary tract infection', <https://doi.org/10.17026/LS/DJBMQJ>, DANS Data Station Life Sciences, V2
- M.C.J. Huisman (Leiden University), 2023, '3D models accompanying Master thesis: Investigating the limits of 3D scanning and modeling', <https://doi.org/10.17026/AR/MNLSIC>, DANS Data Station Archaeology, V1
- C. Weijs (Gemeente Dordrecht), 2023, 'Dordrecht Voorstraat 142-144', <https://doi.org/10.17026/AR/9WRJ4X>, DANS Data Station Archaeology, V1



Bezoek de DANS Data Stations via dans.knaw.nl/nl/data-stations/ of scan de QR-code.

Klimaatonderzoek in Humanities Cluster

Het KNAW Humanities Cluster (HuC) heeft drie projecten toegekend gekregen uit het Onderzoeksfonds KNAW-instituten. De onderzoeksprojecten zijn het resultaat van het initiatief van KNAW HuC om samen te werken op het gebied van Environmental Humanities. Het initiatief is mede tot stand gekomen naar aanleiding van het Klimaatonderzoek Initiatief Nederland (KIN) van NWO en KNAW. Het gaat om de onderzoeksprojecten 'Historische energietransities en AI-toepassingen' van het Huygens Instituut, 'Energietransitie en maatschappelijk draagvlak: de anti-kernenergiebeweging als historische casus' van het IISG en 'Botanical records through a social lens: Exploring bias and perception in historical citizen science' van het Meertens Instituut en Nederlands Instituut voor Ecologie. (MJ)

<https://www.knaw.nl/nieuws/zes-projecten-krijgen-toekenning-uit-het-onderzoeksfonds-knaw-instituten>

Odeuropa's Olfactory Storytelling Toolkit

Vanuit het project Odeuropa is een gids ontwikkeld voor het werken met geuren in musea en erfgoedinstellingen. De gids is gemaakt voor conservatoren, curatoren, docenten, reisleiders, museumdirecteuren, archivariissen, bibliothecarissen en alle anderen die geïnteresseerd zijn om met geuren te werken in een erfgoedcontext. Deze gids is een hulpmiddel dat de basis biedt voor het gebruik van geur als verteltechniek. Het biedt methoden – van begin tot eind – om een olfactorisch verhaal van het (fysieke) collectie-item naar de betrokkenheid van bezoekers te brengen. Met tips voor het opbouwen van een sterk reukverhaal en 'best practices' voor het gebruik van geuren. (MJ)

<https://odeuropa.eu/the-olfactory-storytelling-toolkit/>

Nationale Monitor Geldzorgen

Het afgelopen jaar zagen we overal flinke prijsstijgingen. En dat zorgde bij veel mensen in een kwetsbare financiële situatie voor problemen. Om hier meer inzicht in te krijgen ontwikkelde de Universiteit Leiden en Wijzer in geldzaken de Nationale Monitor Geldzaken, uitgevoerd door Centerdata. Voor deze monitor worden elk kwartaal vragenlijsten voorgelegd aan een representatieve groep Nederlanders (LISS panelleden). Zo brengen onderzoekers de financiële situatie van Nederlanders in kaart en achterhalen ze waar Nederlanders zich zorgen over maken. Hierbij kijken ze ook naar financiële weerbaarheid. Meer weten over de Nationale Monitor Geldzaken? Resultaten van nieuwe metingen worden elk kwartaal bekendgemaakt op de website van Wijzer in geldzaken. De datasets worden ook gepubliceerd in het LISS data archive. (MV)

www.wijzeringeldzaken.nl/Nationale-Monitor-Geldzorgen/

150 uur videomateriaal over buitenlands beleid van Nederland

‘Het geheugen is een ingang naar het verleden’

In het oral history-project ‘Diplomatieke getuigenissen’ zijn 94 videogesprekken met voormalige Nederlandse diplomaten ontsloten. Het geeft een uniek inkijkje in het buitenlands beleid. *Mathilde Jansen*

Het buitenlands beleid van Nederland is al lange tijd een onderzoeksthema van het Huygens Instituut (KNAW HuC). Maar niet eerder werd een collectie met videomateriaal opgeleverd, die ook nog eens heel goed doorzoekbaar is. Het is een primeur in Nederland, vertelt projectleider Cees Heere. “Op deze manier wilden we een collectie creëren die het papieren archief kan duiden én aanvullen.” Heere nam zelf alle 94 interviews af die in de database te vinden zijn. Van elke video is ook een transcript gemaakt. Door de synchronisatie van tekst en video is de zoekfunctie erg dynamisch.

Zonder ruggespraak

Dat een collectie met mondelinge bronnen kennishiaten op kan vullen, staat voor de onderzoeker buiten kijf, vooral als het gaat om dit specifieke onderwerp. “Er zijn diplomaten die in de Koude Oorlog achter het IJzeren Gordijn gestationeerd waren, die continu gevolgd werden door de veiligheidsdiensten van het land waar ze verbleven. Daar is dus nooit iets over opgeschreven. Wij hebben ze gevraagd: hoe ging dat in zijn werk op de ambassade, welke maatregelen werden er genomen om ervoor te zorgen dat je je werk



Amerikaanse Minister van Buitenlandse Zaken, dr. Henry Kissinger brengt bezoek aan Nederland; persconferentie, Kissinger en Van der Stoep, 1976. Credit Bert Verhoeff / Anefo

goed kon doen. Overigens waren alle mensen die ik heb gesproken gepensioneerd, zodat ze konden praten zonder ruggespraak.” De diplomaten die geïnterviewd zijn, waren werkzaam in het buitenland van de late jaren 1950 tot de vroegere jaren 2020. Er is rekening gehouden met een goede spreiding van onderwerpen, zoals geopolitiek, ontwikkelings-samenwerking, Europese integratie, belangenbehartiging van bedrijven en consulaire diensten. Via-via en met hulp van het Ministerie van Buitenlandse Zaken kwam Heere bij de juiste personen terecht. Dat de geïnterviewden terugkijken op een periode in het verleden is volgens de onderzoeker een voor-

deel: “Achteraf kun je vaak beter bepalen wat destijds doorslaggevende factoren zijn geweest.”

Collectieve dialoog

Heere is zich ervan bewust dat het geheugen van de getuigen ook gekleurd is door het heden. “Dat we deze gesprekken hebben gevoerd tegen de achtergrond van de Oekraïne-oorlog is bijvoorbeeld van invloed geweest op de manier waarop mensen zich momenten in de geschiedenis herinneren. Zoals het uiteenvallen van de Sovjet-Unie. Er zijn meerdere diplomaten uit dat gebied die zich gesprekken met willekeurige voorbijgangers

herinneren die hen vertelden: dit communistische systeem kan niet eeuwig duren. Doordat die voorspelling is uitgekomen, herinneren mensen zich dit. Dus het geheugen is niet neutraal, het is een ingang naar het verleden. Maar dat geldt ook voor het geschreven archief, dat is ook niet interpretatievrij. Er bestaat immers ook zoiets als een papieren werkelijkheid. Historici hebben de taak al die verschillende bronnen samen te brengen. Je zoekt continu naar een collectieve dialoog. Zo wordt het beeld van het verleden steeds completer.”

diplomatieke-getuigenissen.huygens.knaw.nl/

Samen werken aan knelpunten

FAIR Data en Open Science gestimuleerd

Het TDCC SSH lanceert gemeenschappelijke projecten voor meer herbruikbare onderzoeksdata en software. *Ricarda Braukmann*

Het thematisch Digital Competence Centre (TDCC) voor Social Sciences and Humanities (SSH) is een initiatief dat een jaar geleden is opgezet om investeringen te faciliteren die ten goede komen aan FAIR en Open Science projecten in het SSH-domein.

Synergetisch netwerk

Dit eerste jaar stond in het teken van het verder definiëren van het doel en de uitgangspunten en de ontwikkeling van de bestuursstructuur van het TDCC SSH. “We zien het TDCC als een netwerk van mensen die het academische systeem willen veranderen, verbeteren wellicht, in de context van Open Science. Door

partnerschappen aan te gaan en samen te werken aan de belangrijkste thema's waar het SSH-domein mee te maken heeft. Activisten van de Open Science-beweging als het ware.” vertelt Nicole Emmenegger, Netwerk Manager van het TDCC SSH.

Knelpunten aanpakken

Het TDCC SSH biedt financiële steun aan projecten die de zogenoemde 'bottlenecks' van het vakgebied aanpakken, zoals vastgelegd in het stappenplan van het TDCC. Diverse onderzoeksteams, bestaande uit zowel onderzoekers als ondersteunend personeel, hebben reeds gedetailleerde projectvoorstellen ontwikkeld om de geïdentificeerde knelpunten aan te pakken. Deze voorstellen zijn beschikbaar op de website van het TDCC. De verwachting is dat de eerste projecten van start zullen gaan in het voorjaar van 2024 en een looptijd kunnen hebben van maxi-

maal 12 maanden. De innovatieve aanpak van het starten met een idee en dit vervolgens ontwikkelen op basis van feedback van de gemeenschap heeft reeds opmerkelijke resultaten opgeleverd. Volgens Nils Arlinghaus, Community Coördinator voor het TDCC SSH, is er een concreet voorbeeld van een team dat na het delen van hun idee binnen het netwerk werd benaderd door een collega van een andere universiteit. “Deze collega bleek lokaal aan een zeer vergelijkbaar project te werken, maar de resultaten waren nog niet openbaar gemaakt. Deze anekdote illustreert duidelijk het voordeel van een aanpak die gefocust is op de gemeenschap, waardoor duplicatie wordt voorkomen en silo's worden doorbroken.”

Uniek proces

Naast de kleinere bottleneck-projecten is er in november 2023 een

tweede financieringsronde geopend voor wat 'challenge-projecten' wordt genoemd. Dit betreft grotere projecten met een maximale duur van 24 maanden. Net zoals bij de bottleneck-projecten gaat het hier om samenwerkingsprojecten die complexe vraagstukken aanpakken die de grenzen van individuele organisaties overstijgen. Volgens Arlinghaus is “het aanvraagproces voor deze oproep enigszins uniek. Voordat er een uitgebreid voorstel wordt ingediend bij de NWO, worden projectideeën eerst gezamenlijk ontwikkeld via het op de gemeenschap gerichte proces. Deze open en inclusieve aanpak bevordert co-creatie en dient ervoor te zorgen dat de uiteindelijke projecten beter aansluiten bij de werkelijke behoeften van de communities binnen het SSH-landschap.”

Grootse dag

Het eerste grote evenement van het

TDCC wordt op 27 februari in Utrecht georganiseerd. De drie TDCC's (SSH, Life Sciences & Health, en Natural & Engineering Sciences) en het Landelijk Coördinatiepunt Research Data Management (LCRDM) nodigen de gemeenschap uit voor een bijzondere dag met als thema ‘Teaming Up Across Domains’. Emmenegger en Arlinghaus verzekeren dat het geen doorsnee dag met formele presentaties zal zijn. “Met dit evenement streven we naar levendige discussies en de vorming van groepen die gedurende het komende jaar gezamenlijk domein-overstijgende uitdagingen willen aangaan.”

Meer informatie over het evenement: tdcc.nl/evenementen/teaming-up-across-domains/ Meer informatie over TDCC-SSH: tdcc.nl/about-tdcc/ssh/

PI en co-PI over de nieuwe onderzoeksinfrastructuur SSHOC-NL

‘Een goed ontworpen infrastructuur merk je niet eens op’

Een gezamenlijke onderzoeksinfrastructuur voor de sociale én geesteswetenschappen. Daar werken CLARIAH en ODISSEI aan in SSHOC-NL.

Erica Renckens

Als nationale onderzoeksinfrastructuren zijn CLARIAH en ODISSEI inmiddels gevestigde namen in de Nederlandse academische wereld. De eerste maakt geesteswetenschappelijk onderzoek met digitale data mogelijk, de laatste is primair gericht op de sociale wetenschappen. In SSHOC-NL (Social Science and Humanities Open Cloud for the Netherlands) slaan beide partijen de handen ineen voor de ontwikkeling van een brede, grootschalige en duurzame digitale infrastructuur die beide vakgebieden bedient. Samen met vijftien partners kregen CLARIAH en ODISSEI afgelopen jaar voor SSHOC-NL €15,2 miljoen financiering toegekend uit de Nationale Roadmap Groot-schalige Wetenschappelijke Infrastructuur van NWO.

Tom Emery (ODISSEI, Principal Investigator (PI)) en Richard Zijdemann (CLARIAH, co-PI) vertellen over het belang en de plannen van de nieuwe infrastructuur.

Waarom zijn jullie deze samenwerking aangegaan?

Emery: “De sociale en de geesteswetenschappen delen veel van dezelfde uitdagingen. Die kunnen we het best aangaan op grote schaal, gedragen door een zo breed mogelijke pool aan expertise. Ontwikkelingen in de kunstmatige intelligentie en computationele methoden hebben bijvoorbeeld een grote invloed op hoe we ons onderzoek uitvoeren en brengen veel vragen met zich mee op het gebied van data privacy, intellectueel eigendomsrecht, niet-transparante werkwijzen en systematische bias.”

Zijdemann: “Daarnaast beschikken de sociale



Tom Emery

en geesteswetenschappen over verschillende bronnen voor gemeenschappelijke onderzoeksvragen. Tot dusver zijn deze bronnen vooral geïsoleerd bestudeerd, aangezien ze zijn beschreven volgens de standaarden van het eigen vakgebied. Door samen te werken in SSHOC-NL kunnen we bronnen generiek aanbieden en de hele SSH-gemeenschap bedienen.”

Welke nieuwe mogelijkheden brengt deze samenwerking met zich mee?

Emery: “Een heel brede range aan mogelijkheden. De grenzen tussen gestructureerde en ongestructureerde data vervagen steeds meer door inspanningen op het gebied van data-harmonisatie. Dat maakt heel nieuwe onderzoekslijnen mogelijk tussen verschillende disciplines. We kunnen bijvoorbeeld nu de analyse van rijke content als kranten en afbeeldingen combineren en ze op grote schaal linken aan meer gestructureerde data als surveys en bevolkingsregisters. Dat geeft

onderzoekers makkelijker toegang tot een breder spectrum aan data voor het beantwoorden van hun onderzoeksvragen. Zo kunnen onderzoekers interdisciplinaire, complexe vragen beantwoorden, die eerder niet konden worden geadresseerd.”

Wat zijn op korte termijn de concrete plannen van SSHOC-NL?

Emery: “Die draaien vooral om het integreren van bestaande services van CLARIAH en ODISSEI, zoals het oplijnen van datacatalogi en microservices, en het gezamenlijk aanbieden van een omgeving waarin veilig met data gewerkt kan worden. Zowel CLARIAH als ODISSEI hebben inmiddels een goed fundament en er zijn veel links tussen de beide gemeenschappen, wat betekent dat we vanaf het eerste jaar aan de gang kunnen met het leveren van verbeterde diensten. Een goed voorbeeld daarvan is de SANE-omgeving waarin onderzoekers veilig toegang krijgen tot een brede range aan data die eerst buiten bereik lagen. Deze dienst zal vanaf februari 2024 beschikbaar zijn.”

Wat hopen jullie uiteindelijk met deze samenwerking te bereiken?

Emery: “We hopen vooral spannende, nieuwe samenwerkingen tussen de sociale en geesteswetenschappen te zien, waardoor innovatieve nieuwe ideeën en theorieën breed hun weg kunnen vinden. Zo veel vragen uit de sociale wetenschappen worden gedeeld met, of komen zelfs oorspronkelijk van onze collega’s uit de geesteswetenschappen. Met een multidimensionaal en interdisciplinair perspectief kunnen we vast en zeker onze kennis verbreden en verdiepen.”

Zijdemann: “Een ander belangrijk voordeel van deze samenwerking ligt in het delen van een infrastructuur, die kostbaar is om te onderhouden. Zo beschikt CLARIAH op dit moment niet over een beveiligde analyse-omgeving, bijvoorbeeld voor documenten met IPR, en daarvoor is samenwerken met SANE erg gunstig. Voor de sociale wetenschappen zal het grote aantal nieuwe SANE-gebruikers weer

voordelig uitkomen bij het verwerven van financiering van het toekomstige onderhoud van SANE.”

Wat gaan onderzoekers in de sociale en de geesteswetenschappen merken van deze samenwerking?

Emery: “Een goede infrastructuur neemt barrières weg en als die goed is ontworpen en gebouwd, merk je niet eens dat hij er is. Onze hoop is dat over een jaar of vijf een jonge onderzoeker toegang heeft tot uiteenlopende data uit het hele SSH-domein en deze kan linken en analyseren en dit alles heel natuurlijk en eenvoudig vindt werken. Ik hoop echt dat de volgende generatie de diensten die we nu ontwikkelen als vanzelfsprekend ziet en zich er niet eens van bewust is ze te gebruiken.”

Zijdemann: “Door SSHOC-NL zullen onderzoekers aan beide zijden van het SSH-domein meer inzicht krijgen in elkaars werk, wat interdisciplinaire samenwerkingen katalyseert. Databronnen zullen meer geïntegreerd en makkelijker doorgrond kunnen worden, tools worden generieker, waardoor ze net zo goed kunnen worden ingezet voor de analyse van Twitter-data als van advertenties in 19e-eeuwse kranten.”

Hoe gaan jullie de gebruikers, de onderzoekers, betrekken in dit project?

Emery: “Bij het bouwen van een infrastructuur gaat het om hardware, software en ‘peopleware’. Dat laatste is misschien wel het belangrijkste. Je moet onderzoeksgemeenschappen samenbrengen om te ontdekken welke diensten nodig zijn, technici stimuleren om deze te onderhouden en onderzoekers trainen om ze te gebruiken. De ontwikkeling van de infrastructuur is een zeer mensgericht proces en we horen dan ook altijd graag van mensen wat ze nodig hebben en wat ze willen. We nodigen dan ook iedereen uit om contact met ons op te nemen, maar gaan zelf ook actief de boer op.”

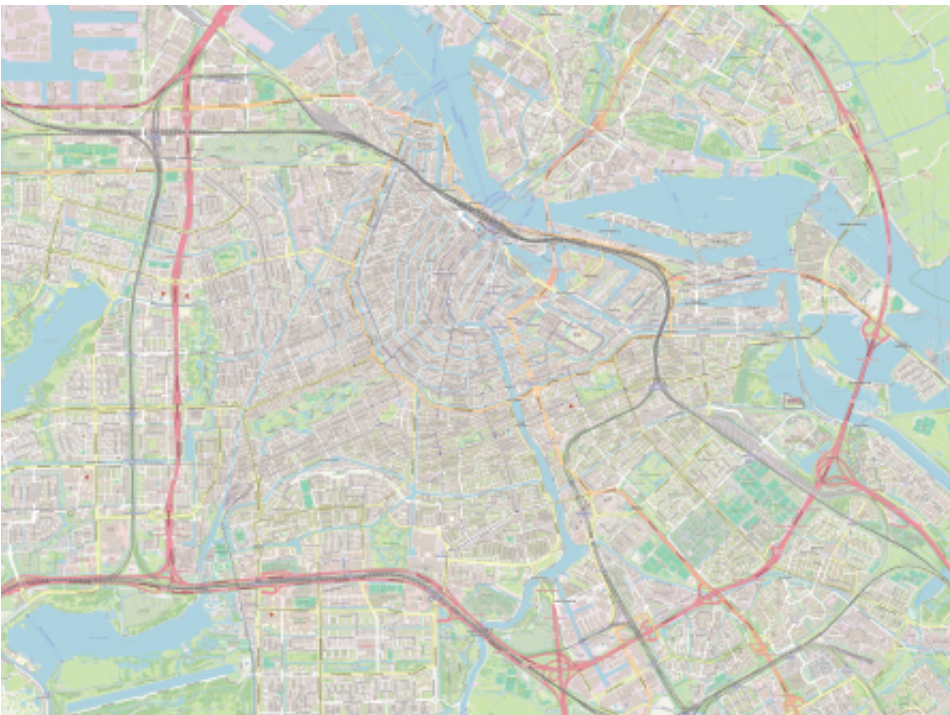
Zijdemann: “Ik sluit me volledig bij Tom aan. Eén van vijf werkpakketten in SSHOC-NL is gericht op de ‘peopleware’: communicatie, disseminatie, trainingsmaterialen en integratie met bestaande netwerken. Als onderzoeker moet je echter altijd de afweging maken tussen iets snel doen en iets goed doen. Het kost tijd om te leren werken met een nieuwe of uitgebreide infrastructuur. Als onderzoekscollectief is het logisch om te investeren in een gedeelde infrastructuur, maar voor een onderzoeker lijkt de initiële investering misschien niet de moeite waard voor ‘een enkel paper’. We horen heel graag van onderzoekers met zulke bedenkingen, of van mensen die misschien betere oplossingen hebben. Want dat biedt veel voordelen. Als infrastructuur kunnen we nog altijd veel leren over wat onderzoekers nodig hebben. Bovendien proberen we de initiële leercurve zo vlak mogelijk te maken. Uiteindelijk zullen bronnen onderling steeds meer met elkaar verbonden zijn, waardoor het alleen via gedeelde infrastructuren mogelijk zal zijn om echt verrijkte onderzoeksdata en -methoden te bieden.”

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

<https://odissei-data.nl/nl/>



Richard Zijdemann



Kaart van Amsterdam uit 1832 en 2023. Bron: HisGIS.nl/OpenStraatMap contributors

HisGIS plakt historische kaarten op het moderne landschap

Kadasterkaarten tonen landschap van toen

Op HisGIS.nl kun je terug in de tijd navigeren naar de Napoleontische kadasterkaarten uit 1832. Met nieuwe subsidie wordt de komende jaren gewerkt aan een nationale infrastructuur. Mathilde Jansen

Het HisGIS-project startte ooit bij de Fryske Akademy, maar is inmiddels ondergebracht bij het KNAW Humanities Cluster. Mede-projectleider Rombert Stapel (IISG) vertelt hoe de databank sinds de jaren 90 steeds groter is geworden: na de gedigitaliseerde kadasterkaarten uit Friesland volgden kaarten uit steeds meer provincies. Maar door de langdurige groeipe-

riode sloten al die bestanden technisch niet meer op elkaar aan. “We kunnen nu wel mooie plaatjes laten zien, maar HisGIS is nog niet bruikbaar als nationale infrastructuur”, zegt Stapel. “Hier gaan we de komende jaren verandering in brengen.”

Tekenen met de hand

In het HisGIS-project worden oude kadasterkaarten gekoppeld aan referentiepunten in het moderne landschap, vertelt Stapel. “Bijvoorbeeld een slootje dat er nu nog loopt of een hoek van een pand dat er nog staat. Als je dat voor heel Nederland doet, kun je die historische kaarten op de moderne kaarten plakken.” Onderzoekers noemen dit ook wel ‘georefereren’. De volgende stap is het zogenaamde ‘vectoriseren’: daarbij worden perceelsgrenzen van de oude kadasterkaarten overgetekend. Met de hand, omdat computerprogramma’s

hier nog niet precies genoeg voor zijn. Voor het handmatig overtekenen van de kaarten is een grote groep vrijwilligers aangetrokken. Wat tot slot ook nog gedigitaliseerd moet worden uit het kadaster zijn de tabellen die gegevens bevatten over de percelen. Voor historici bevatten deze tabellen ontzettend veel informatie, vertelt Stapel enthousiast. “Hoe het perceel gebruikt werd, wie de eigenaar was, wat zijn of haar beroep was. Er zitten lokale verkoopprijzen in van bijvoorbeeld graan, maar ook lokale loonprijzen, pachtprizen, transportkosten. Dat soort zaken werden allemaal nauwkeurig bijgehouden. Al die data geven een waardebeoordeling van een perceel.”

Het landschap van vroeger

Met die informatie kun je allerlei sociaalhistorisch onderzoek doen, maar vooralsnog wordt de website nog te weinig geraadpleegd

door historici. Dat komt doordat de technische infrastructuur het nog niet altijd goed mogelijk maakt om systematisch onderzoek te doen, maar daar komt de komende jaren verandering in. Met een subsidie vanuit de Rijksdienst voor Cultureel Erfgoed en het Ministerie van OCW zal een grote slag gemaakt worden in het uniformeren van de data.

De kaarten worden momenteel wel volop gebruikt door andere doelgroepen, vertelt Stapel, zoals archeologische adviesbureaus en beleidsbureaus. “Bij nieuwbouw moet altijd onderzocht worden of ergens archeologische resten te verwachten zijn. Daarvoor worden deze kaarten veel gebruikt. Of om te achterhalen hoe het landschap er vroeger uitzag. Bij natuurherstel is het nuttig om te weten wat voor natuur er in de 19e-eeuwse ondergrond nog te vinden is.”

Bovendien worden de kaarten gebruikt door mensen die geïnteresseerd zijn in hun lokale geschiedenis. Zo kun je bijvoorbeeld opzoeken wie er vroeger in jouw huis woonde. Stapel gebruikt de kaarten wel alvast voor zijn eigen onderzoek. “Ik ben mediëvist, en dus gespecialiseerd in oudere perioden dan de 19e eeuw. Maar de kadasterkaarten geven in feite de situatie van Nederland vóór de industrialisatie. Daardoor kun je zelfs het middeleeuwse landschap erin terugvinden.”

<https://hisgis.nl/>

De ODISSEI-codebibliotheek

Code delen voor betere wetenschap

Bij het publiceren van wetenschappelijk onderzoek, zoals een artikel in een tijdschrift, is het een goed gebruik om ook de analytische code te delen. Angelica Maineri

Daarmee worden alle stappen bedoeld die nodig zijn om van de ruwe gegevens tot de gepubliceerde resultaten te komen, inclusief alle gegevensverwerking en analytische stappen. Het gaat dan bijvoorbeeld om een syntaxisbestand van Stata of SPSS of een codeerschema van een kwalitatief onderzoek. Het verza-

melen en veilig opslaan van dit materiaal, bijvoorbeeld als onderdeel van gegevenspakketten, wordt steeds vaker vereist door instellingen en tijdschriften, dus waarom niet de extra stap nemen en de gegevens openbaar beschikbaar maken?

Voordelen van code delen

Misschien zijn mensen terughoudend met het delen van hun code, omdat ze bang zijn dat er fouten worden gevonden. Dit kan iedereen overkomen, maar alleen door transparant te zijn is het mogelijk om fouten te corrigeren en te verbeteren. In feite zitten er veel voordelen aan het delen van code. Het verhoogt bijvoorbeeld de betrouwbaarheid van onderzoek omdat anderen alle

stappen kunnen bekijken en – als het ook mogelijk is om de onderliggende gegevens te delen – de resultaten kunnen repliceren. Het delen van code stelt de gemeenschap ook in staat om van elkaar te leren, doordat onderzoekers oplossingen voor gezamenlijke problemen opnieuw kunnen toepassen.

Hoe code delen

Er zijn verschillende onlineplatforms die geschikt zijn voor het delen van code. GitHub is erg populair voor versiebeheer, maar garandeert geen langdurige opslag. Daarvoor kunnen beter platforms als Open Science Framework (OSF) en Zenodo worden gebruikt, waar ook een directe verbinding met GitHub-

repository’s kan worden gemaakt. Het is niet nodig om ook de onderliggende ruwe gegevens op te nemen, aangezien gegevens vaak wel toegankelijk zijn, maar gebruikers deze niet mogen herdistribueren. Daarom is het bij het delen van code belangrijk om documentatie toe te voegen waarin niet alleen wordt uitgelegd hoe de code gebruikt moet worden, maar ook welke gegevens nodig zijn en hoe deze toegankelijk zijn. Voordat code wordt gedeeld, is het ook belangrijk om te controleren of er geen privacygevoelige informatie in de code zelf staat.

ODISSEI-codebibliotheek

Zelfs als gegevens goed gearchiveerd zijn, kan het voor andere on-

derzoekers lastig zijn om ze te vinden. Om code die gebruikmaakt van de ODISSEI-infrastructuur beter vindbaar te maken, is ODISSEI begonnen met het opbouwen van een bibliotheek van analytische code die is gebruikt in studies die gebruikmaken van LISS-data of CBS-microdata. Voor elk project wordt, naast de link naar de code, de link naar de paper en, indien beschikbaar, een DOI getoond die direct verwijst naar de gebruikte dataset.

Wilt u bijdragen of heeft u ideeën voor verbetering? Stuur een e-mail of plaats uw bijdrage op het GitHub-project!

angelicamaineri.github.io/ODISSEI-code-library/

Een alternatieve brainstormmethode

Design thinking in kwalitatief onderzoek

Design thinking is een methode die gebruikt wordt om complexe problemen op te lossen door een creatieve manier van denken. Origineel werd het voornamelijk gebruikt voor het ontwerpen van producten, maar steeds meer en meer wordt het ook toegepast om ook in de onderzoekswereld (complexe) problemen op te lossen. Mara Verheijen



Brainstormmethode. Bron: Pixabay

In design thinking wordt in stappen naar één of meerdere oplossingen voor een maatschappelijk vraagstuk toe gewerkt. Het is een creatievere methode dan een klassiek interview of dialoogtafel en bij design thinking staat de mens centraal; de ervaringen en behoeften van de doelgroep zijn cruciaal bij het ontwikkelen van een oplossing. Hoe begin je met de design thinking methode?

Inleven

De design thinking methode bestaat uit vijf stappen. Het inleven vanuit het perspectief van de doelgroep is de eerste fase. Het doel is om ervoor te zorgen dat je altijd alle verschillende perspectieven van de doelgroep boven tafel krijgt. Tijdens dit onderdeel van de dialoogtafel kun-

nen verschillende methodieken worden gebruikt, zoals bijvoorbeeld een empathy map. Een empathy map is een visualisatie van de gedachtes, gevoelens, en gedragingen van een bepaalde doelgroep en wordt gebruikt voor het inleven in de wereld van deze doelgroep. Zonder te weten wat er omgaat in anderen (gevoelens, gedachten, etc.) bestaat het risico iets te ontwikkelen dat helemaal niet aansluit bij wat mensen nodig hebben.

Definiëren

In deze stap definieer je samen met je doelgroep wat de kern is van het probleem.

Verbeelden

Dit is de meest creatieve fase in het proces en hét moment waarin je ideeën ophaalt. In deze

fase worden namelijk zo veel mogelijk oplossingen gegenereerd voor het probleem. Het gaat hierbij niet om het vinden van de perfecte oplossing, maar om zo veel mogelijk verschillende oplossingsrichtingen te bedenken. Het kan dus eigenlijk gezien worden als een creatieve brainstorm.

Tijdens deze creatieve brainstorm worden de deelnemers uitgedaagd om “out of the box” te denken en op andere manieren naar het probleem te kijken zodat er innovatieve oplossingen voor het probleem ontstaan. Hiervoor wordt bijvoorbeeld de wishful thinking methode gebruikt om de ideale situatie te beschrijven; stel dat er geen regels en geen obstakels zijn, wat zou je dan doen? Een andere methode is de slechtst-mogelijk-idee methode. Het doel hierbij is om de meest slechte ideeën

te bedenken. Zo’n idee biedt waarschijnlijk wel een oplossing voor het probleem maar is ‘slecht’, met oog op bijvoorbeeld ethische aspecten, uitvoerbaarheid en financieel vlak. Door deze methode toe te passen en te bespreken wat deze ideeën zó slecht maakt kan men elementen destilleren die juist wél bruikbaar zijn.

Prototype

Hier gaat het om het creëren van een snelle en eerste versie van de oplossing, die past bij de doelgroep/eindgebruiker. Het doel in deze fase is om snel een ruwe versie van een prototype te maken, bijvoorbeeld met legoblokjes of een tekening op papier. Door snel een prototype te maken ben je niet eindeloos tijd kwijt aan het perfectioneren van een oplossing die uiteindelijk toch niet blijkt te werken.

Testen en optimaliseren

In deze fase wordt getest hoe de doelgroep de gekozen oplossing ervaart, wat werkt wel en wat werkt niet? Vervolgens wordt dit ge-optimaliseerd.

De vijf stappen hoeven niet noodzakelijkerwijs allemaal doorlopen te worden, dit is afhankelijk van het onderzoeksdoel. Het onderzoeksdoel kan ook zijn om zo veel mogelijk oplossingen op te halen, maar deze nog niet uit te voeren. In dat geval stop je na stap drie.

Al met al is design thinking een creatieve methode die ruimte biedt voor innovatieve ideeën en out of the box oplossingen. Wil je meer weten over design thinking als onderzoeksmethode, neem dan contact op met Centerdata:

www.centerdata.nl

COLUMN

Laat duizend bloemen bloeien?

In bijna elk overleg waar ik aan deelneem, komt de vraag langs: wie doet eigenlijk wat op het terrein van onderzoeksdata? Soms is het ook echt een verzuchting: hoe komen we nu echt verder met de ontwikkeling van diensten voor onderzoeksdata als er zo veel initiatieven tegelijk lopen? Zeker, de digitalisering van de wetenschap en de open science agenda’s van instellingen en financiers (NWO, Europese Commissie) leiden ook voor de alfa- en gammawetenschappen tot een scala aan programma’s op het terrein van (FAIR) onderzoeksdata. En dat is natuurlijk erg fijn op een belangrijk terrein waar nog zó veel te ontwikkelen valt. Maar wie nu een visie en plannen wil ontwikkelen, of zelfs concrete projectdoelen wil formuleren, heeft er best een klus aan om onontgonnen terrein te identificeren en dubbel werk te voorkomen. Want er wordt volop geëxperimenteerd en uitgeprobeerd: elke editie van E-Data staat er vol mee, en terecht! Dit betekent dat tools,

ordeningen en onderdelen van digitale infrastructures dicht bij wetenschap ontwikkeld worden; in veel gevallen ook in teamverband waarin diverse kennis en ervaring samenkomt met die van de gebruiker, de onderzoeker. En al die experimenten zijn ook nodig omdat er nog geen gezamenlijk beeld is van de meest optimale infrastructures en diensten en van de weg(en) daarnaartoe: Wat kan bijvoorbeeld het beste ontwikkeld worden binnen wetenschapsdisciplines en wat misschien via opschalen van generieke oplossingen? Menno Rasch gaf daar in zijn column in dit blad, oktober 2023, ook een mooi voorbeeld van: is het misschien ook zinvol diensten te ontwikkelen langs de lijn van de rol van

data in onderzoek en dus de omgang met data? En natuurlijk maken AI en ML vooruitkijken nog ingewikkelder.

Zelf denk ik dat deze veelheid van initiatieven en experimenten op het terrein van onderzoeksdata erbij hoort, nu. Want



Anja Smit

we staan pas aan het begin van een ontwikkeling. In een stadium zoals dat waarin geleerden elkaar nog brieven schreven om hun kennis uit te wisselen en er nog geen tijdschriften bestonden om dat proces te structureren. Waarin brieven nog geen wetenschappelijke artikelen waren en marktpartijen die kennisdeling nog niet konden organiseren, harmoniseren en uiteindelijk monopoliseren. Kortom, het stadium waarin de oplossingen, hoe onvoldragen ook, nog gestuurd worden binnen het publieke domein en niet door (commerciële) financiële business cases.

Daarmee is deze fase, waarin we ook binnen het publieke domein oplossingen kunnen zoe-

ken, dicht bij de wetenschap, mijns inziens van grote waarde. De diversiteit in biedt ruimte voor keuzes die optimaal aansluiten bij wat de wetenschap nodig heeft om te innoveren. En opdrachtgeverschap binnen het publieke domein zorgt voor publieke waarden als fundament. Tegelijk is natuurlijk de roep om consolidatie, samenwerking langs duidelijke lijnen absoluut zinvol en ook nodig. Immers, we kunnen onmogelijk alle experimentele oplossingen in leven houden en ook de beschikbare middelen zinvol besteden. Laten we dus blijven proberen zicht te krijgen op wie nu eigenlijk wat doet op het terrein van onderzoeksdata én onze eigen goede ideeën met anderen blijven delen, ofwel: samen innoveren.

Anja Smit is directeur van DANS, Instituut van KNAW en NWO. DANS is expertisecentrum en dienstverlener op het terrein van onderzoeksdata. Anja geeft de column door aan Alastair Dunning.