

Van analog naar digitaal bruikbaar WO2-onderzoeksarchief

TRIADO maakt archieven doorzoekbaar

Hoe maak je van kilometers ongestructureerd archiefmateriaal bruikbare digitale onderzoeksdata? Naar die methodiek is het project Tribunaalarchieven als Digitale Onderzoeksfaciliteit (TRIADO) op zoek.
Netwerk Oorlogsbronnen

Het meest geraadpleegde Tweede Wereldoorlogarchief van Nederland, het Centraal Archief Bijzondere Rechtspleging (CABR), dient als casus voor TRIADO. Op dit moment is het CABR alleen doorzoekbaar via een database, op naam van verdachten, en zijn de stukken niet digitaal raadpleegbaar. TRIADO, dat mogelijk is gemaakt dankzij het KNAW-Onderzoeksfonds, ontsluit en verkrijgt voor het eerst op een geautomatiseerde wijze dit rijke WO2-archief digitaal en op documentniveau. Hierdoor wordt het mogelijk om gegevens over tientallen personen, gebeurtenissen en plaatsen in dossiers te doorzoeken. En dus om grootschalig vergelijkend onderzoek te doen.

Digitale technieken

Het Netwerk Oorlogsbronnen, Nationaal Archief, Huygens ING en NIOD Instituut voor Oorlogs-, Holocaust- en Genocidestudies zijn in 2017 van start gegaan met het generieke deel van het project. Hierin wordt



Een fysiek CABR-dossier wordt met TRIADO digitaal doorzoekbaar gemaakt.
foto Nationaal Archief

onderzocht welke digitale technieken gebruikt kunnen worden om hybride archiefmateriaal digitaal te ontsluiten. Begin 2018 is een steekproef uit het CABR gedigitaliseerd en doorzoekbaar gemaakt (ge-OCR'd). Met speciale software voor Named Entity Recognition en automatische datumextractie wordt vervolgens informatie over onder meer personen, locaties en datumaanduidingen vastgesteld. De uit de documenten gehaalde gegevens over wie, wat, waar en wanneer worden gekoppeld aan referentiedata uit onder meer de

WO2-thesaurus en de Basisregistraties Adres- en Gebouwen. Marielle Scherer, data-manager Huygens ING: "Deze verrijking-fase zorgt voor structurering van de data en een betere doorzoekbaarheid van de dossiers uit de steekproef".

Nieuwe vragen

Het projectteam is enthousiast. Anne Gorter, bij het Nationaal Archief projectleider nadere toegankelijkheid CABR: "Door deel te nemen aan dit project bereiden we ons voor op

de toekomst en op klantvragen die gaan komen". Edwin Klijn, projectleider TRIADO en programmamanager Netwerk Oorlogsbronnen: "Ik kan haast niet wachten om te zien welke nieuwe vragen we kunnen stellen aan dit archief en hoe digitale methodieken ons daarbij kunnen helpen".

Data laten spreken

Voor het tweede, onderzoeksdeel van TRIADO formuleert onderzoeker Ismee Tames van het NIOD met een wetenschappelijke adviescommissie verschillende kwantitatief-statistische onderzoeksvragen. Tames: "Denk aan vragen als: 'Wie werd wanneer door Nederlandse collaborateurs gearresteerd? Hoe verhoudt dit zich tot de ontwikkeling van het Duitse vervolgingsbeleid? Wie waren de collaborateurs en verzetslieden eigenlijk: zien we veranderingen in bijvoorbeeld leeftijd, gender of regio?' Dat we nu de data tot ons laten spreken in plaats van een van tevoren opgestelde lijst met namen terug te zoeken, is een veelbelovende stap voorwaarts".

bit.ly/TRIADOproject

Het Netwerk Oorlogsbronnen wordt gefaciliteerd door het NIOD Instituut voor Oorlogs-, Holocaust- en Genocidestudies en gefinancierd door het Ministerie van VWS, Vfonds en VSBfonds.

SINDS KORT BESCHIKBAAR

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

CentERdata

• God in Nederland

Verschuivingen in de christelijke godsdienst, de godsdienstigheid van Nederlanders en de rol van de kerken in het publieke debat, daar wil het onderzoek 'God in Nederland' zicht op krijgen. De meest recente meting is in juni 2015 afgenomen in het LISS panel door Joris Kregting (KASKI, Radboud Universiteit). Voor de eerste keer via een websurvey werd de respondenten gevraagd hoe men in het eigen leven tegen religie en spiritualiteit aankijkt en op welke wijze men daar mee bezig is. Het onderzoek wordt sinds 1966 iedere tien jaar gehouden en is in boekvorm gepubliceerd. Het databestand God in the Netherlands is beschikbaar via LISS Data Archive. lissdata.nl

Ook sinds kort beschikbaar:

Studies LISS panel

- CentERdata, april 2017 - mei 2017, Work and Schooling - Wave 10
- CentERdata, juni 2017 - juli 2017, Economic Situation: Housing - Wave 10
- Hubner, S.; Soest, A. van, juli 2017, Time Use and Consumption - Wave 5
- Scheele, D.; Fouarge, D., juni 2017 - juli 2017, Dutch Skill Survey (NSS) 2017
- Voet, J. van der, mei 2017, ABW Waste disposal



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

DANS

• Data nieuw aardewerkonderzoek beschikbaar

Het project Data Atlas van Byzantijnse en Osmaanse materiële cultuur: het archiveren van Middeleeuwse en Post-Middeleeuwse archeologische veldwerkgegevens uit het oostelijk Middellandse Zeegebied (600-2000), Phase 1 (Vroom, Dr. J.A.C., Faculty of Archaeology, Leiden University, 2017) toont de resultaten van aardewerkonderzoek vanuit opgravingen en surveys in het oostelijke Middellandse Zeegebied. Mede hierdoor is nu meer bekend over de typochronologie van aardewerk uit de periode 600 na Christus tot heden binnen het onderzoeksgebied en is meer kennis verzameld over onder meer de relatie tussen productie, distributie en gebruik van aardewerk in de vier stedelijke centra Athene, Butrint, Efeze en Tarsus. Het onderzoek is onderdeel van het NWO-VIDI Project (276-61-003) 'Material Culture, Consumption and Social Change: New Perspectives for Understanding the Eastern Mediterranean during Byzantine and Ottoman Times' (2010-2015) en ontving een KDP-subsidie vanuit DANS voor het ontsluiten van de data.

DOI: 10.17026/dans-zuq-ewst

Ook sinds kort beschikbaar:

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

- Huijgen, T.G. (Research Centre for Education and the Labour Market – ROA) (2017): HBO-Monitor 2016. DANS. DOI: 10.17026/dans-z5w-4yhk
- Chamberlain, B.P. (Imperial College London) (2017): Real-Time Community Detection in Full Social Networks on a Laptop. DANS. DOI: 10.17026/dans-2bc-4qgc
- Stoof, R (2017): Studentenonderzoek 2017. DANS. DOI: 10.17026/dans-xph-jdnu
- Kohnstamm Instituut, Universiteit van Amsterdam (2016): Pre-COOL tweejarigen-cohort vijfde meting 2014-2015. DANS. DOI: 10.17026/dans-xjb-8wj



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



Keramische vondsten van de opgravingen.
credits N. Gail. oeaw.ac.at/oeai/forschung/bauforschung/ephesos-ayasuluk-forschungen

Wie wint de Nederlandse Dataprijs 2018?

Komend najaar wordt weer de Nederlandse Dataprijs uitgereikt. Een prijs voor een onderzoeker of onderzoeksgroep die extra bijdraagt aan de wetenschap door onderzoeksdata beschikbaar te maken voor aanvullend of nieuw onderzoek. Bij de vorige editie in 2016 sprak de jury over 'allemaal mooie voorbeelden van het toegankelijk maken en delen van onderzoeksdata'. De organisatie van de Nederlandse Dataprijs is in handen van Research Data Netherlands, een samenwerkingsverband tussen 4TU.Centre for Research Data, DANS en SURFsara. Binnenkort staat meer informatie over de Dataprijzen 2018 op de website van RDNL. (HB)

researchdata.nl

Update voor OpenSoNaR+

De webinterface van OpenSoNaR+, het referentiecensus voor geschreven en gesproken Nederlands, heeft begin dit jaar een update ondergaan. De 540 miljoen woorden aan allerhande teksten en ruim 900 uur aan spraakopnames in het corpus zijn nu toegankelijk voor iedereen met een CLARIN-account. Bovendien is de zoekfunctie flink verbeterd, met uitgebreidere zoekmogelijkheden en snellere resultaten tot gevolg. Zo kunnen gebruikers de zoekresultaten nu groeperen op contextwoorden of gebruikmaken van een random sample als steekproef. (ER) opensonar.clarin.inl.nl/