

NWO haalt de teugels weer strakker aan

‘Die data zijn niet van jou’

De regels van Nederlands grootste onderzoeksfinancier voor open data en open access worden steeds strenger.

Warna Oosterbaan

NWO wordt mede-eigenaar van de data die in een gesubsidieerd onderzoek worden geproduceerd. De instelling kan die data dan weer aan andere onderzoekers ter beschikking stellen. Dit staat in de nieuwe Regeling Subsidieaanvragen van NWO, die op 1 mei van kracht werd en die voor alle wetenschappelijke disciplines geldt. De onderzoeksfinancier wil op die manier het hergebruik van data bevorderen.

Volgens Ron Dekker, directeur Instituten bij NWO leidt hergebruik tot een beter rendement op wetenschappelijke resultaten, en voorkom je zo dat voor het produceren van onderzoeksgegevens “twee keer betaald moet worden”. NWO verstrekt per jaar ruim 550 miljoen euro aan onderzoekssubsidies.

Bij het beoordelen van onderzoeksvorstellen zal er voortaan ook strenger op gelet worden of

de data die de onderzoekers willen verzamelen, misschien al zijn te vinden in de dataset van een ander onderzoek. “Als dat zo is, dan is het einde oefening”, zei Dekker op de ‘Open Onderzoeks-Data Dag’ die op 18 mei door Surffoundation was georganiseerd.

Dekker maakte op die dag bekend dat NWO naast het invoeren van deze regels een campagne gaat starten waarin onderzoekers erop wordt gewezen dat hun onderzoek met publieke middelen wordt gefinancierd. “En we gaan ze ervan bewust maken: die data zijn niet van jou. Ze zijn van de instelling en van NWO. En daarom moeten die data goed en toegankelijk worden gedocumenteerd en opgeslagen”.

Steun in de rug

Peter Doorn, directeur van data-archief DANS zegt in een reactie de nieuwe bepaling “een steun in de rug van DANS” te vinden. “Wij kunnen NWO helpen door data-overeenkomsten met de subsidieontvangers te sluiten. Dat doen we al voor de investeringssubsidies bij de gebieden Geesteswetenschappen en Maatschappij- en Gedragwetenschappen, maar we breiden dat met liefde uit naar de andere vakgebieden. Daarbij

zullen we het credo hanteren: ‘open als het kan, beschermd als het moet’.”

In de nieuwe voorwaarden zijn ook de regels die betrekking hebben op de vrije toegang tot wetenschappelijke publicaties (*open access*) verder aangescherpt. Auteursrechten mogen niet worden overgedragen als open access daardoor onmogelijk wordt. Een andere regel stelt dat NWO ter bevordering van de kennisoverdracht “nadere voorwaarden kan stellen aan de auteursrechten”.

In een toelichting laat de voorlichter van NWO weten dat hier gedacht kan worden aan een boek voor een breder publiek over een bepaald onderwerp. “NWO dient dan te beschikken over de nodige licenties om de door haar gesubsidieerde publicaties te mogen gebruiken voor het uit te geven boek.” Impliceert ‘gebruiken’ dan letterlijke overname? NWO: “Ja, overname van het artikel in zijn oorspronkelijke staat.” En als de uitgever van het oorspronkelijke artikel dat niet toestaat? NWO: “Als er uit een programma een artikel voortkomt dat ter publicatie wordt aangeboden, dan moet op dat moment al duidelijk zijn dat er op het gebruik van de inhoud een claim ligt. Uiteraard heeft de uitgever dan de mogelijkheid om af te zien van publicatie.”

Zoekmachine vindt relevante WO2-verwijzingen in Handelingen der Staten Generaal

Dat doet denken aan de oorlog

Verwijzen naar de Tweede Wereldoorlog is een beladen kwestie in het politieke debat. Het lijkt wel of politici dat steeds vaker doen. Maar is dat ook zo? Erica Renckens

Het NIOD Instituut voor Oorlogs-, Holocaust- en Genocidestudies ontwikkelt samen met de Universiteit van Amsterdam een intelligente zoekmachine die automatisch relevante verwijzingen naar de Tweede Wereldoorlog vindt in de Handelingen der Staten Generaal. Zo kunnen onderzoekers precies zien wanneer een politicus in welke context en met welk doel de oorlog aanhaalde.

“Tot nu toe was dat niet goed mogelijk”, zegt Hinke Piersma, onderzoekster bij het NIOD. Een normale zoekmachine gaat alleen uit van het ingetypte trefwoord en geeft daarvoor ook veel resultaten die niet relevant zijn. Bijvoorbeeld dat het brood sinds de Tweede Wereldoorlog veel duurder is geworden. “Deze zoekmachine kan dat wel, omdat hij alle woorden uit het document meeneemt in zijn analyse. Zo kunnen we inzicht verschaffen in de wijze waarop de Tweede Wereldoorlog doorwerkt in de Nederlandse politieke arena.” Het zoekprogramma wordt ontwikkeld in het kader van CLARIN, de Europese digitale onderzoeksomgeving voor geesteswetenschappers.

“Onze zoekmachine kijkt naar alle



Mei 1945. Een lid van de Binnenlandse Strijdkrachten gebruikt een foto van Hitler als schietschijf foto Beeldbank WO2, Verzetsmuseum

woorden die eerder door training met de Tweede Wereldoorlog geassocieerd zijn”, aldus Piersma.

Wel of niet relevant

In deze training hebben historici voor een klein gedeelte van de Handelingen der Staten Generaal aangegeven of de documenten wel of

niet verwijzingen naar de oorlog bevatten. Op basis hiervan is een onderwerpspecifiek model getraind dat automatisch van de overige documenten kan bepalen of zij wel of niet relevant zijn voor het onderwerp Tweede Wereldoorlog. “Bij deze analyse gebruiken we alle woorden die tijdens de training in de

“Als het hebben van een dubbele nationaliteit de weg zou openen naar verraad en disloyaliteit, dan stel ik daar tegenover dat iemand als Mussert, voor zover ik weet, niet de Duitse nationaliteit had.”

Door de zoekmachine gevonden citaat van Mohammed Rabbae (GroenLinks) tijdens een debat op 22 februari 1995.

als relevant geannoteerde documenten voorkwamen én de woorden uit de niet-relevante documenten”, legt Piersma uit. “Een woord dat in beide categorieën ongeveer even vaak voorkomt beschouwen we als ‘neutraal’. Komt een woord onevenredig vaak voor in één van beide collecties dan zien we dat als een sterke aanwijzing om het document als ‘relevant’ of juist als ‘niet-relevant’ te categoriseren.”

Een veelgehoord geluid is dat het hedendaagse politieke debat veruwt, met meer verwijzingen naar de Tweede Wereldoorlog. “Deze vergelijkingen zijn er echter altijd geweest”, zegt Piersma. “Maar door de hele naoorlogse periode in het onderzoek te betrekken is het wellicht mogelijk om trends en patronen hierin te laten zien.”



E-DATA & RESEARCH

Jaargang 6 | 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: DANS, CentERdata, CBS, CLARIN-NL, Huygens ING, Koninklijke Bibliotheek en de Vereniging voor Geschiedenis en Informatica

INHOUD

3
Strektalen met Gabmap in beeld gebracht

4
De ideale digitale bibliotheek in tien stellingen

4
500 veteranen leggen ziel en zaligheid bloot

5
20.000 krantenartikelen over kunst in de database

6
Tapes en ponskaarten in het computermuseum

7
Hollandse graven hebben hun oorkonden op het net

7
Jon Orwant van Google: ‘We willen iets terugdoen’

8
Statistiek en sushi op Summer School

COLOFON

Uitgever: Stichting Uitgeverij E-data & Research Den Haag

Redactieadres: Postbus 93067, 2509 AB Den Haag, 070-3446484 edata@dans.knaw.nl

Redactie: Inge Angevaere, Eric Balster, Heidi Berkhout, Ronald van der Bie, Peter Boot, Warna Oosterbaan (hoofd-/eindredacteur), Thijs Hermesen, Erica Renckens, Jetske van der Schaaf

Redactiesecretariaat: Lucas Pasteruning, Jetske van der Schaaf

Aan dit nummer werkten mee: Jan Burgers, Ingrid Dillo, Marieke van Erp, Maria Fleischmann, Barend van der Meulen, Hilje Oosterbaan Martinius, Leo van Velzen, Steamwork Graphics

Opmaak: Colette Sloots, Haarlem

Productie: Amsterdam University Press

Druk: Ten Brink, Meppel

Oplage: 9100.

ISSN: 1872-0374

E-data & Research is online te raadplegen op www.edata.nl. Toezending is kosteloos aan relaties van de stakeholders en op verzoek aan studenten in de alfa- en gamma-richtingen.