

De cruciale rol van data-archieven

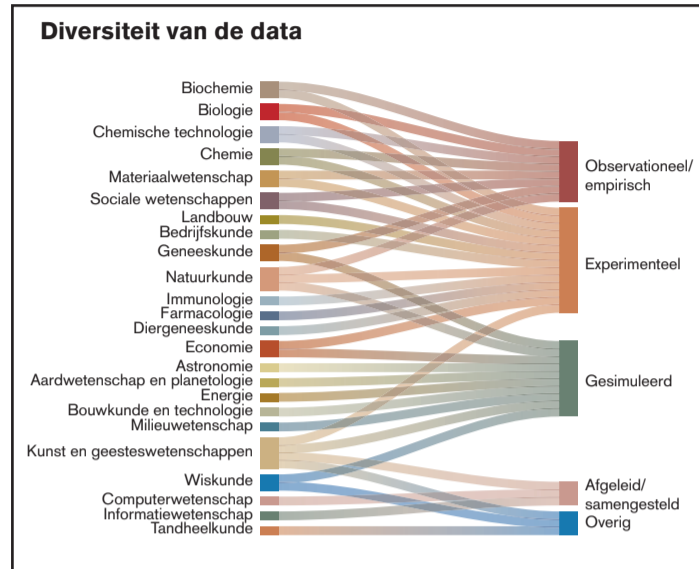
Hoe zoeken gebruikers naar data van andere onderzoekers?

Hergebruik van door anderen gecreëerde data kan een veelbelovende onderzoeksstrategie zijn, maar er is weinig bekend over de manier waarop gebruikers naar dergelijke data zoeken. *Kathleen Gregory*

Data-archieven spelen een belangrijke rol bij het ontsluiten van zogenoemde secundaire data. Toch is er maar weinig bekend over de manier waarop gebruikers naar deze data zoeken. Wat zijn hun motieven? Waarvoor hebben ze de data nodig? Wat voor soort data zoeken ze?

1.677 respondenten

Onlangs hebben onderzoekers van DANS, de Universiteit Maastricht en de Universiteit van Amsterdam de resultaten van het grootste bekende onderzoek naar deze vragen gepubliceerd in het kader van een project gefinancierd door NWO creative Industries Grant in samenwerking met Elsevier. 1.677 respondenten uit 105 landen, uit verschillende disciplines, met verschillende rollen en in verschillende fasen in hun academische loopbaan, vulden de vragenlijst in. De uitkomsten van het onderzoek, gepubliceerd in het artikel



Uit 'Diversiteit van de data' (linker grafiek) blijkt de diversiteit van de data die onderzoekers uit verschillende disciplines nodig hebben. Iets meer dan de helft van de respondenten geeft aan data van buiten de eigen discipline nodig te hebben.

'Gebruik van data' (rechter grafiek) toont het kerngebruik van de data door de respondenten; 71% zei data te gebruiken als basis voor een nieuwe studie. De grafiek laat zien dat het datagebruik wordt bepaald door discipline domeinen en methoden; zo worden domeinen die doorgaans gebruikmaken van computeronderzoeksmethoden, geassocieerd met het gebruik van data voor model- en algoritme-input. Voor beide afbeeldingen geldt: statistisch significante associaties gedetecteerd met Bonferroni-correctie voor gelijktijdige paarsgewijze marginale onafhankelijkheid; $n = 1677$; significantieniveau: $p < 0,05$. Credits: Kathleen Gregory en Steamwork Graphics

Lost or found? Discovering data needed for research (Gregory, K., Groth, P., Scharnhorst, A. & Wyatt, S. (2020). Harvard Data Science Review. DOI: 10.1162/99608f92.e38165eb), geven inzicht in de databehoeften van de respondenten, de bronnen en strategieën die ze gebruiken om data te vinden en de criteria die

ze hanteren om die data te evalueren.

De onderzoeksgegevens zijn vrij toegankelijk via EASY: Gregory, K.M (Data Archive and Networked Services) (2020): Data Discovery and Reuse Practices in Research. Dit artikel is een vertaling van het

Engelstalige artikel van Kathleen Gregory, M.A., MSLIS, PhD-kandidaat, werkzaam bij het project Re-SEARCH bij Research & Innovatie binnen DANS.

DOI: 10.17026/dans-xsw-kkeq

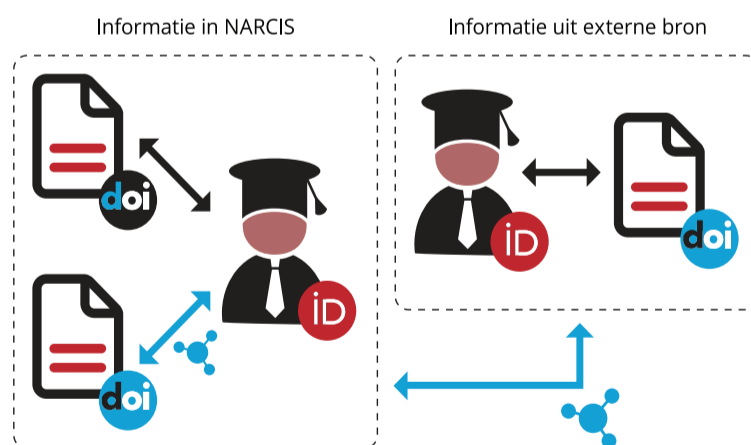
DANS verbindt en verrijkt beschikbare informatie in NARCIS

PID Graph van de Nederlandse wetenschap

Benieuwd naar informatie over de Nederlandse wetenschap? De portal NARCIS.nl biedt steeds meer informatie door de automatische koppeling van persistent identifiers in een PID Graph.

Ricarda Braukmann

Een persistent identifier (PID) is vergelijkbaar met een Burgerservicenummer voor digitale objecten; het is een uniek en blijvend nummer waaraan allerlei informatie is gekoppeld. Het gebruik van PIDs is van groot belang om op een betrouwbare manier te kunnen verwijzen naar het juiste artikel of de juiste persoon of organisatie. Aan veel wetenschappelijke informatie is dan ook zo'n PID toegekend. Door deze PIDs en door de koppeling van een PID aan een andere PID, verbetert de vindbaarheid, toegankelijkheid en navigatie van onderzoeksgegevens. Het koppelen



PID Graph maakt relaties zichtbaar. In dit voorbeeld zijn twee papers van 1 onderzoeker via NARCIS beschikbaar. Van 1 paper is bekend dat het hoort bij deze onderzoeker. Omdat NARCIS eigen informatie linkt met informatie uit externe bronnen (bijvoorbeeld ORCID), wordt ook de link met het tweede paper bekend. De NARCIS PID Graph haalt deze externe informatie op en integreert het in NARCIS, waardoor de link tussen het tweede paper en de onderzoeker ook zichtbaar wordt, gevisualiseerd met het blauwe PID Graph symbool. Credits: Ricarda Braukmann

van PIDs kan automatisch doordat PIDs ook leesbaar zijn voor computers. Het Europese project FREYA onderstreept de kracht van PIDs en bouwt aan een infrastructuur voor PIDs als essentieel onderdeel van de

European Open Science Cloud (EOSC).

De nationale portal NARCIS.nl biedt informatie over (open access) publicaties, datasets, onderzoeksprojecten, wetenschappers en onder-

zoeksinstellingen. Onlangs is de informatie in NARCIS verrijkt met behulp van bestaande en nieuwe PIDs. Zo zijn onderzoeksorganisaties nu eenduidig identificeerbaar in NARCIS doordat ze nu de nieuwe PID research organisation identifier (ROR) bevatten.

Relaties zichtbaar

Door PIDs met elkaar te verbinden in een PID Graph kunnen meer relaties zichtbaar gemaakt worden in NARCIS. Zo vergelijkt NARCIS bijvoorbeeld de informatie van het portaal van ORCID (een PID voor onderzoekers) met de beschikbare informatie in NARCIS en linkt publicaties aan de onderzoeker als deze nog niet als auteur was herkend. Op dezelfde manier is een link gemaakt met unpaywall.org, waardoor NARCIS nu ook informatie geeft over open access versies van bepaalde wetenschappelijke artikelen. Een laatste voorbeeld is de beschikbaarheid via NARCIS van de informatie van altmetric.com, een webservice over de online impact van papers of datasets op bijvoorbeeld twitter of facebook. De NAR-

CIS PID Graph maakt het ook mogelijk om resultaten te verzamelen van bepaalde onderzoeksprojecten. In een pilot project hebben het Donders Instituut, de Radboud Universiteit en DANS publicaties en datasets uit het NWO-gefinancierde project Language in Interaction aan elkaar gekoppeld.

De PID Graph biedt dus veel mogelijkheden om de informatie in NARCIS te verrijken. Hiervoor is het wel van belang dat PIDs door onderzoekers en instellingen worden gebruikt en meegeleverd in de metadata.

narcis.nl
project-freya.eu

AGENDA

Door het coronavirus zijn veel evenementen geannuleerd of is het onduidelijk of evenementen doorgaan. Veel organisaties bieden digitale alternatieven, minstens zo interessant. We adviseren u om online te zoeken naar actuele informatie.