

MEEGEMAAKT

Internationale samenwerking en het paardentram-paradigma

Suzan van Dijk

Aangespoord door artikelen in *e-data&research*, en aangelokt door de digitaliserings- en internationaliseringsgelden, willen ook geesteswetenschappers wel eens grootschalig samenwerken. Voor bepaalde onderwerpen is het trouwens noodzakelijk. Toch doen zich op dat front problemen voor, waaraan men aanvankelijk niet gedacht had. Naar aanleiding van een brede internationale samenwerking op het gebied van vrouwenliteratuur ga ik hier op enkele aspecten in – rekening houdend met wat Hugo Brandt Corstius tien jaar geleden schreef: ‘Het is met vrouwenliteratuur als met de paardentram’. Deels zijn de te signaleren problemen misschien terug te brengen tot dit gegeven, want: ‘U denkt misschien bij “paardentram” aan een tram die door paarden wordt getrokken. Maar een paardentram is ook een tram waarin paarden zich van de wei naar de rennen laten rijden. Of waar een olijk gemeentebestuur paardenlijven op heeft laten schilderen...’. De materie is, kortom, ingewikkeld en kan van verschillende kanten worden benaderd.

Een aanvraag voor zo'n meestal Europese samenwerking wordt veelal voorbereid door een groep van vijf à tien collega's. In onderling overleg schrijven zij een onderzoeksprogramma dat in het ideale geval inderdaad voor financiering in aanmerking komt. In dat overleg houdt men rekening met mogelijke, en soms bekende behoeften van andere, later uit te nodigen collega's. Gepoogd wordt om recht te doen aan de te verwachten diversiteit van perspectieven. Zo was het althans in ons geval.

Inmiddels is onze samenwerking volop op gang gekomen, en is men in twintig landen druk doende om de eerste werkbijeenkomst voor te bereiden. De gedachte was dat eenieder daarbij enthousiast gebruik zou maken van enkele reeds bestaande, volgens ons heel redelijk functionerende

online tools. In de praktijk blijkt het echter anders te werken. Ten eerste doet zich het bekende probleem voor van de tegenover ict nog steeds zeer aarzende geesteswetenschapper, en ten tweede hadden wij zelf misschien het 'paardentram-paradigma' nadrukkelijker moeten expliciteren.

‘Waarom stappen onderzoekers in een digitaliseringsoperatie?’, vroegen wij ons regelmatig af toen wij mailtjes begonnen te ontvangen van collega's uit diverse landen, waarin zij berichtten dat (1) bepaalde gegevens niet konden worden ingevoerd, (2) *idem* voor gegevens uit een bepaald land, (3) *idem* voor gegevens die afhankelijk lijken van bepaalde andere, nog niet ingevoerde, gegevens, en vooral (4) dat zij hun wachtwoord waren vergeten. ‘Waarom bekijken onderzoekers niet eerst de reeds aanwezige content?’, was een andere vraag, toen bleek dat heel wat al ingevoerde informatie ten tweeden male aan de database werd toegevoegd. Op onze vraag wáár precies de Instructies tekort schieten kwam geen antwoord.

De problemen zijn niet alleen van technische aard. Sommige collega's, die alleszins in staat lijken om de data in te voeren, doen dat niet omdat ze vrezen dat hun bronnenmateriaal door anderen voortijdig (vóór de publicatie van hun conclusies) ontdekt wordt. Ze gaan er kennelijk van uit dat dit materiaal slechts voor één inter-

pretatie vatbaar is. *Quod non* ongetwijfeld, ook al betreft het vrouwenliteratuur.

Wat onze eigen paardentram-benadering van vrouwenliteratuur betreft, concludeerden wij dat wij wellicht wat optimistisch zijn geweest. Waarschijnlijk ligt er een voorbereidende taak voor geesteswetenschappers die rond een enigszins breed onderwerp een samenwerking op touw willen zetten. Van het paardentram-effect is niet iedereen zich bewust, en daardoor ook niet van de noodzaak om de eigen oogkleppen eens kritisch te bekijken.

Omdat wij nog aan het begin van onze samenwerking staan houden wij de moed erin: gezien de verklaringen van 'nieuwe' collega's die willen toetreden, en ook toelichten waarom zij dat willen, is daar alle reden toe. Het belang dat wordt gehecht aan vrouwenliteratuur en aan de bestudering ervan is per land zeer verschillend: het is niet alleen literair-historisch maar ook, zeker in Oostelijk Europa, politiek. Dit spoort op zichzelf weer aan tot het realiseren van de samenwerking. De 'technische' problemen zullen, in ons geval, dan ook wel worden overwonnen.

Suzan van Dijk is verbonden aan de faculteit Geesteswetenschappen van de universiteit Utrecht en gastonderzoeker bij het Huygens Instituut



Gelezen

Bojan Todosijevic, Kees Aarts, Harry van der Kaap: *Dutch Parliamentary Election Studies, Data Source Book 1971-2006*; Den Haag, DANS, 2010; *DANS Data Guide 7*, ISBN 978 94 90531 04 1

General elections for the Second Chamber of Parliament constitute one of the most important elements of democracy in the Netherlands. Insight into the backgrounds, motives, goals and behaviour of the Dutch voters is therefore essential for understanding and appraising the functioning of democracy. Since 1971, the Dutch political science community has conducted large scale, nationwide election surveys around every parliamentary election. All resulting datasets have been extensively documented and archived, and placed at other researchers' disposal through DANS. This Data Guide provides the documentation of an integrated dataset based on the Dutch Parliamentary Election Studies of 1971-



2006. It thus provides directly accessible information on change and stability in electoral behaviour and political orientations in the Netherlands. The data

are concentrated on the more frequently asked questions of these Dutch Parliamentary Election Studies.

Michiel Hildebrand: *End-user support for access to heterogeneous linked data SIKS; Dutch Research School for Innovation And Knowledge Systems, 2010; SIKS Dissertation Series no. 2010-22*

The openness of the Web allows data from different sources to be combined, enabling information access in ways unforeseen by the original providers. The rise of linked data adds a new dimension to this information revolution. Using Semantic Web representation languages, structured information becomes more easily acceptable by machines, enabling emergent technologies to reuse, integrate and process information. The foundations for the Web of data are now in place and institutions are starting to share linked data. In this thesis, Michiel

Column

Robbert Dijkgraaf

Dataprijs

Een oude regel in de wetenschap zegt dat je je onderzoeksgegevens altijd beschikbaar moet houden voor wie je conclusies zou willen checken. Eigenlijk is die ene regel al genoeg om elke discussie over het delen van data snel te beslechten. Je zou daar nog aan toe kunnen voegen dat het delen van data niet alleen goed is om elkaar te controleren, maar ook om elkaar te stimuleren. Dan heb je al twee sterke argumenten.



Maar de tijden zijn veranderd, en die verandering brengt in een stijgend tempo nieuwe overwegingen mee. Overwegingen die allemaal in dezelfde richting wijzen. Om te beginnen de digitalisering, die het behalve wenselijk ook veel beter mogelijk heeft gemaakt om gegevens uit te wisselen. De oprichting van het Steinmetzarchief in de jaren zestig van de vorige eeuw laat zien dat de sociale wetenschappers dat voordeel al wilden verzilveren in het vroeg elektronische tijdperk. Maar toen het elektronische tijdperk eenmaal was aangebroken, kwam er in veel disciplines een wereld aan nieuwe mogelijkheden bij: om data te delen en bij elkaar te voegen, en om veel grotere hoeveelheden data te verzamelen en verwerken dan tot dan toe denkbaar was geweest. Om dat te illustreren hoef ik de Geneefse deeltjesversneller maar te noemen, die niet alleen de samenwerking van duizenden onderzoekers vereist, maar ook zelf voor een groot deel de technologie heeft ontwikkeld die deze samenwerking mogelijk maakt.

Het elektronische tijdperk verandert de wetenschap niet alleen in kwantitatief maar ook in kwalitatief opzicht. Zo zie je in de humaniora, waar de KNAW juist dit jaar een impuls geeft aan de toepassing van ict in het onderzoek, dat die toepassingen nu al leiden tot hele nieuwe vormen van analyse en samenwerking.

Intussen wordt de ons omringende maatschappij hectischer, ingewikkelder, veeleisender en kritischer. Dat betekent dat er meer druk op de wetenschap komt te staan, dat er grotere problemen op ons bordje worden gelegd en dat we kritischer worden beoordeeld - op onze conclusies door iedereen, getuige ook de discussies over de al of niet beschikbare data van het klimaatpanel IPCC – en door de overheden op de kosteneffectiviteit van ons onderzoek. Ook dat alles wijst in de richting van meer en, waar mogelijk, interdisciplinair samenwerken, meer krachten bundelen en meer openheid. Dus data bundelen en data beschikbaar maken.

Het elektronische tijdperk heeft ons niet alleen maar goeds gebracht. Door de snelle ontwikkeling van hard- en software blijken digitale gegevens heel wat vergankelijker te zijn (in letterlijke zin) dan de data die we vroeger op papier bewaarden. Er is dus veel meer aandacht voor duurzaamheid nodig om te zorgen dat de vruchten van digitaal datadelen zich ook werkelijk laten plukken.

Moet ik meer zeggen om uit te leggen waarom ik op 18 juni graag zelf de DANS Dataprijs in Amsterdam wil overhandigen aan de alfa- of gamma-onderzoeksgroep waarvan de beste, meest toegankelijke en duurzaamste dataset is genomineerd?

Prof. dr. Robbert Dijkgraaf is president van de Koninklijke Nederlandse Akademie van Wetenschappen

Hildebrand investigates access to heterogeneous linked data from an end-user perspective. Based on case studies in the cultural heritage domain he explores how domain experts can benefit from semantically-rich linked data. With the development and evaluation of prototype systems he demonstrates how a number of specific tasks can be supported. Hildebrand shows that generic solutions can be successfully applied to linked data, by using the semantic relations in the data combined with careful configuration of the search algorithms and interface components.

Tony Hey, Stewart Tansley, Kristin Tolle (Eds.): *The fourth paradigm, data intensive scientific discovery*; Redmond

Washington, Microsoft research, 2009; ISBN 978 0 9825442 0 4

Increasingly, scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets. The speed at which any given scientific discipline advances will depend on how well its researchers collaborate with one another, and with technologists, in areas of eScience such as databases, workflow management, visualization, and cloud computing technologies. In this publication, the collection of essays expands on the vision of pioneering computer scientist Jim Gray for a new, fourth paradigm of discovery based on data-intensive science, and offers insights into how it can be fully realized.