

# Verlag werkconferentie Biografisch Portaal op 10 december 2010

*Rik Hoekstra*

Op 10 december 2010 organiseerde het Biografisch Portaal een workshop voor ontwikkelaars en wetenschappers over vraag hoe het Biografisch Portaal geschikt is te maken voor gebruik als wetenschappelijk instrument. Deze vraag heeft verschillende kanten:

- hoe vinden we aansluiting bij de wetenschap
- welke technische mogelijkheden zijn er
- hoe wortelt het portaal in de praktijk van het (menswetenschappelijk) onderzoek

De technische aangelegenheden kwamen 's ochtends aan de orde en de relatie tot de praktijk van onderzoeksprojecten 's middags.

Kees Mandemakers gaf de dag een achtergrond vanuit het perspectief van de Historische Steekproef Nederland (HSN op <http://www.iisg.nl/hsn/indexnl.html>), een grootschalig en succesvol onderzoeksproject van het IISG. Hij legde sterk de nadruk op de verbondenheid met het onderzoek; de steekproef is ooit opgezet vanuit een aantal onderzoekers die met soortgelijke vragen zaten en de krachten bundelden. Andere factoren die hebben bijgedragen aan het succes van de HSN waren de samenwerkingsprojecten en de internationale oriëntatie.

Maarten Marx en Jacco van Ossenbrugge gingen beiden in op aspecten van de technische mogelijkheden vanuit hun eigen achtergrond in het verbinden van materiaal uit verschillende bronnen. Hun insteek was verschillend. Marx koppelt in zijn Political Mashup (<http://www.politicalmashup.nl>) uiteenlopende bronnen (data en teksten) op een impliciete manier, waarbij hij vooral gebruik maakt van zoektechnologie, terwijl Ossenbruggen werkt aan het semantisch web. Hij maakt gebruik van expliciete metadata als personen, plaatsen en trefwoorden die door en bij de data zelf worden gepubliceerd en vervolgens kunnen worden gecombineerd.

Hoewel hun benadering wat verschilt, hadden ze beiden gelijklopende aanbevelingen voor een grotere bruikbaarheid van het Biografisch Portaal voor anderen. Het is van het grootste belang expliciet, consistent en duurzaam gegevens te publiceren op het net, zodat ze door anderen gebruikt kunnen worden. Gebruik leidt tot hergebruik en zo volgt er cumulatief een zwaan-klief-aan effect. Doordat je gegevens zichtbaar zijn en een rol spelen in wisselende contexten kan iedereen van elkaars inspanningen gebruik maken en profiteren. Hierdoor ontstaan conglomeraten van data die meer zijn dan de som van de delen. Het is belangrijk de data te publiceren zoals ze zijn en niet te bevreesd te zijn voor de fouten die ze nog bevatten, maar ook om ze niet mooier voor dan ze zijn door hiaten of manco's te verdoezelen. De publicatie van gegevens via een vast schema en onveranderlijk adres is een voorwaarde voor de zichtbaarheid en bruikbaarheid.

De technische mogelijkheden hoe de biografische gegevens van het portaal te gebruiken en hergebruiken kwamen hiermee in de ochtendsessie goed naar voren. De vraag van de

middagsessie was hoe ze concreet in feitelijke onderzoeksprojecten een rol kunnen spelen. Hans de Valk, eminence grise van de bronuitgave, schetste een breed scala aan mogelijkheden die de onderzoeker heeft gekregen met de grootschalige beschikbaarheid van digitale bronnen.

Nieuwe databestanden roepen hun eigen vragen op, vragen die bij het samenstellen niet zijn te voorzien en al helemaal niet in combinatie met andere bronnen. De bronontsluiter moet wel luisteren naar de specifieke vragen van zijn directe gebruikers, maar met het oog op toekomstige gegadigden zijn ontsluiting ook zo generiek mogelijk houden door formele criteria te volgen – hij kent immers de vragen nog niet. Een structurele uitgave zoals het Biografisch Portaal is zo altijd een halffabricaat, dat velen tegemoet komt maar vrijwel altijd nog extra werk en verrijking van de onderzoeker behoeft. Dit bleek ook uit de lezingen van Antheun Janse en Martin Bok, over hun eigen projecten, respectievelijk de Bourgondische adel en Encartico, een project (en database) over zeventiende-eeuwse Amsterdamse schilders en hun markten. Beide projecten zijn geruime tijd geleden begonnen en hebben hun eigen dataverzameling opgebouwd. Het Biografisch Portaal in zijn huidige vorm was, als het toen al had bestaan, weinig bruikbaar geweest, aangezien er niet voldoende informatie over 'hun' personen aanwezig is en de informatie ook niet specifiek genoeg is. Voor concrete samenwerkingsprojecten is dus altijd meer en betere informatie nodig dan wat een algemene voorziening als het Biografisch Portaal kan bieden. Het is waarschijnlijk dat dat wel altijd zo zal blijven; onderzoekers zoeken tenslotte de grenzen op van wat we weten en willen dus altijd meer en beter.

De centrale vraag waarop we antwoord wilden krijgen op deze dag was hoe het Biografisch Portaal als onderzoeksinstrument te ontwikkelen. Voor de zichtbaarheid, voor de positie als referentiepunt en voor de samenwerking met andere projecten is het nodig het Portaal verder open te stellen voor bevraging en *harvesting* door computers. Verder is het is duidelijk dat er meer biografieën in het Portaal moeten, en dat de kwaliteit van de ontsluiting van de opgenomen personen omhoog moet, maar ook dat concrete onderzoeksprojecten altijd specifiekere informatie nodig zullen hebben. Het gaat te ver tegelijk in te willen spelen op alle wensen en vragen, die we bovendien maar zeer ten dele in kaart kunnen brengen. De beste manier hierbij aan te haken lijkt daarom met het Biografisch Portaal direct te gaan participeren in een aantal concrete onderzoeksprojecten. Zo kunnen we de kwaliteit van de biografische data voor een specifieke onderzoeksveld verbeteren en vervolgens de resultaten van die projecten weer incorporeren in het Portaal, voorzover ze biografisch interessant zijn. Daarmee ontstaat een dialoog met het wetenschappelijk veld dat het Biografisch Portaal óók wil bedienen, naast het algemeen geïnteresseerde publiek. De mogelijkheden daartoe zijn zeker aanwezig, zoveel werd duidelijk op 10 december.