

Biografisch onderzoek en Bi(o)g(rafische) data

Brainstorm THATcamp,
Den Haag 14-01-14

Dr. Serge ter Braake et al.

Wat gaan we doen?

- Wat is (g)een goede biografie?
- Wat is big data?
- Beschikbare data
- Mogelijkheden kwantitatief
- Mogelijkheden kwalitatief
- Methodologische valkuilen
- Technologische valkuilen
- Conclusies?

Wat is (g)een goede biografie?

Een goede biografie:

- Gebaseerd op oorspronkelijke, persoonlijke bronnen
- Moet goed leesbaar zijn
- Kritische weging bronnenmateriaal
- Plaats een persoon in zijn tijd/of plaats de tijd in de persoon
- 'Compenseer' voor de focus op één persoon/aspect
- Comparatief (prosopografische elementen)

Geen goede biografie?:

- De rode draad hindert de auteur om een geloofwaardig leven te schrijven
- Focust alleen op de hoofdpersoon

- Nadelen prosopografie:
- Bekijkt een persoon vanuit slechts 1 groep. Oplossing: netwerkonderzoek, idealiter combinatie met biografische elementen

Wat is big data?

- Buzzword. Data is data
- Wanneer is iets groot?
- Doen Humanities niet aan big data?
- Maar: de grens ligt bij wat behapbaar is, of beter met computationele technieken gedaan kan worden.
- Zaken die niet verwerkt kunnen worden door mensen, of behapt, maar die je wel kan ondervragen
- Bruikbaarheid afhankelijk van de vraagstelling
- P. 57 BMGN (Eijnatten/Pieters/Verheul)

Beschikbare data voor biografisch onderzoek

- Historische Steekproef Nederland (tot 1923)
- Biografisch Portaal/Oxford Dictionary of National Biography/ Biographie-portal
- Burgelijke Stand voor zover digitaal (WieWasWie ?)
- Dbpedia
- Delpher
- Archieven.nl
- Mailboxen?
- Interviews via beeld en geluid
- Handelingen van de Eerste en Tweede Kamer der Staten-Generaal
- Twitter?/Internet?/Hyves?
- Crowdsourcen (bv. Bentham project)
- E-books
- Google books
- Je eigen database ? <= weer beschikbaar stellen voor anderen, dan wel koppelen aan het boek
- Etc ...

Mogelijkheden kwantitatief

- Vergelijkingen met anderen \leq telbare karaktertrekken
- Op zoek naar covariantie
- Niet alleen plaatsen in de tijd, maar ook met de tijd \leq markeren events met tijdsaanduiding
- Grote hoeveelheden data kunnen in relatief korte tijd doorzocht worden
- Etc ...

Mogelijkheden kwalitatief

- Netwerken in kaart brengen, ook qua representatie
- Interactieve discussie met de data (proces moet wel inzichtelijk zijn)
- Nieuwe mogelijkheden om visueel onderzoek te koppelen aan tekstonderzoek
- Mogelijkheden te onderzoeken wat mensen **niet** zeggen in vergelijking met anderen
- Historiografisch (hoe werden mensen door de tijd heen beoordeeld?)
- Etc ...

Methodologische valkuilen

- Wat zeggen die big data over een bepaalde kwestie?
- Betrouwbaarheid van bronnen als DBpedia
- Waar is de context? (BGMN, Jeurgens)
- Telkens dezelfde data worden gerehashed voor onderzoek, gaan toekomstige historici nog de archieven in?
- Utopie van alle beschikbare data (BMGN p. 19)
- Etc ...

Technologische valkuilen

- Gegevensverlies door digitalisering (artikel Jeurgens BMGN)
- Het proces van het drukken op de knop tot de uitkomst is vaak niet inzichtelijk voor de historicus
- Ruis in de (soms ongestructureerde) data (Ribbens en Piersma, BMGN)
- Precisie van de uitkomsten laat nog vaak te wensen over
- Etc ...

Conclusies?

- Waar mogelijk, trek vergelijkingen tussen personen
- Let op mogelijkheden en beperkingen van data
- Andere 'paden' van een biografie die niet in de lopende tekst gekomen zijn zichtbaar maken on-line (eigen database, narratieve zijpaden)
- Er moeten cursussen komen voor historici (technisch), zodat ze weten wat er gebeurt na de "druk op de knop"
<= meer doen, minder praten
- DH blijft voor een biografie voorlopig alleen een hulpmiddel; het opent paden, biedt comparatieve mogelijkheden die eerder te arbeidsintensief waren
- Documenteer de keuzes die je gemaakt hebt