

**Rapportage Workshop Geschiedenis en Digitale Historische Tekstanalyse, 4 april 2014,
Vrije Universiteit Amsterdam, Wis- en Natuurkundegebouw, Intertain Lab (S1.11)**

Deelnemers (17):

Voor BiographyNet:

Serge ter Braake
Antske Fokkens
Niels Ockeloen
Piek Vossen

Ingeschreven deelnemers:

Sebastiaan Derks
Thomas D'haeninck
Lisa van Diem
Marie-Christine Engels
Jan Hein Furnée
Dennis Geusebroek
Vincent Kuitenbrouwer
Ferne Maas
Pim Huijnen
Willem Prins
Jacqueline Reeuwijk
Amandine Thiry
Andreas Zangger

Afmeldingen: 6

No show: 5

Introductie

Deze workshop was de vrucht van een idee van computationeel linguïst Antske Fokkens (AF) in het kader van het BiographyNet project. Ze vroeg zich af wat er zou gebeuren als je een groep historici (computationeel) laat nadenken over het beantwoorden van historische vragen aan de hand van de data van het Biografisch Portaal van Nederland. Op die manier kan zij gericht aan de slag met het digitaal analyseerbaar maken van de data en krijgen de deelnemers inzicht in de mogelijkheden en beperkingen (nu en in de toekomst) van digitale historische tekstanalyse. In samenspraak met historicus en projectgenoot Serge ter Braake (StB) en de andere leden van het BiographyNet team is een opzet ontworpen waarmee een vruchtbare uitwisseling van ideeën en kennis werd beoogd binnen de tijd van één middag.

De uitnodiging (bijlage 1) voor deze bijeenkomst is verspreid onder de historici van de VU en de UvA, en via het eScience Center, de KNAW e-Humanities group en LinkedIn (historici.nl). Een hoger dan verwacht aantal mensen schreven zich in (24), van wie uiteindelijk dertien

geïnteresseerden zijn gekomen. Wellicht een teken dat dit onderwerp veel historici interesseert, maar dat het uiteindelijk niet de hoogste prioriteit geniet.

Beschrijving van de middag

Deel 1

De middag begon met een algemene inleiding van StB, AF en Niels Ockeloën over digitale historische tekstanalyse, linked data en RDF. Vervolgens werden vier zogenaamd eenvoudige opdrachten, veelal gericht op het in kaart brengen van groepen personen, verdeeld onder de deelnemers die zich opsplitsen in drie groepjes van drie en een van vier. Het was de bedoeling dat er vooral werd nagedacht over de te volgen methode om aan de hand van de beschikbare data een antwoord te krijgen op een historische vraag. Aangezien de BiographyNet demonstrator nog volop in ontwikkeling is, moesten de deelnemers nog gebruik maken van het zoekformulier van het Biografisch Portaal. Bij het maken van deze opdrachten en de nabespreking kwamen een aantal zaken naar voren die vooral te maken hebben met de huidige beperkingen van het zoekformulier en de 'volledigheid' (voor zover je daar ooit van zou kunnen spreken) van de data. Beperkingen gerelateerd aan het huidige zoekformulier zijn:

1. De eigenschappen waarop gezocht kan worden, zijn beperkt. Het is in het algemeen moeilijk om een groep mensen gebaseerd op bepaalde criteria uit het corpus te halen. Zo kan bijvoorbeeld niet op functie gezocht worden om een groep mensen die hetzelfde beroep hebben uitgeoefend te selecteren.
2. Het zou verder handig zijn als mensen die meerdere eigenschappen met elkaar delen geselecteerd kunnen worden.
3. Het zou handig zijn als lijsten van personen ook direct geïmporteerd kunnen worden (bijvoorbeeld uit Wikipedia: de lijst van gouverneurs-generaal van Nederlands-Indië).
4. Het vrije zoekveld, van fundamenteel belang voor bijvoorbeeld netwerkonderzoek, lijkt niet (optimaal) te werken.
5. Zoekresultaten kunnen alleen alfabetisch worden geordend, terwijl het voor de historicus juist handig zou zijn ze ook chronologisch te kunnen ordenen.
6. Begrippen waarnaar gezocht wordt, zijn niet makkelijk terug te vinden in de teksten. Highlights in de tekst zouden hierbij kunnen helpen.
7. Er zijn geen overzichten van eigenschappen van een persoon (of alleen beperkt).
8. Relaties tussen personen (bijvoorbeeld een stamboom) moeten uit de tekst gehaald kunnen worden. Als dit in RDF staat, zou het toch mogelijk moeten zijn, zulke stambomen te genereren.

De beperkingen die voortkomen uit het (gebrek aan) 'volledigheid' van de data kunnen onder meer aan twee eigenschappen van de data in het Biografisch Portaal gerelateerd worden:

1. Het feit dat alleen overleden mensen in het Portaal zijn opgenomen maakt het beantwoorden van met name vragen die op relatief recente geschiedenis betrekking hebben moeilijk.
2. Ondanks dat het biografisch portaal een ruime collectie aan biografische woordenboeken met verschillende selectiecriteria bevat, zijn niet alle (mogelijke) groepen van mensen even goed vertegenwoordigd.

Deelnemers gaven daarom ook aan dat het voor veel vragen nuttig zou zijn als ook bronnen van buiten het biografisch portaal in het onderzoek meegenomen kunnen worden. Met name links naar Wikipedia zouden hierbij behulpzaam kunnen zijn.

Deel 2

Na een korte pauze gingen de deelnemers in groepjes van een andere samenstelling verder met een zogenaamd moeilijke vraag, vooral gericht op concepten, historiografie en ontwikkelingen. Ook hier kwamen een aantal zaken duidelijk naar voren in de nabespreking. De vragen waren moeilijk, zowel inhoudelijk als om naar de voor de computer leesbare componenten (RDF) te vertalen. We onderscheiden twee soorten uitdagingen, die het beantwoorden van vragen met de huidige tools moeilijk maken. De eerste uitdagingen hebben met name betrekking op historiografie (welke ideeën in de teksten naar voren komen en hoe er over onderwerpen geschreven wordt):

1. Het is moeilijk termen te vinden, die een bepaald oordeel, idee of manier van presenteren goed weergeven. Wat zijn bijvoorbeeld indicatieve termen voor nationalisme? Welke termen duiden op een positief, zacht oordeel over gouverneurs? Welke zijn juist kritisch?
2. Een mogelijke aanpak zou zijn om eerst teksten te analyseren waarvan bekend is dat ze een bepaalde visie of opinie uitdragen (bijvoorbeeld nationalistische teksten) en deze vervolgens te vergelijken met verschillende biografieën.
3. Functionaliteiten die bestaande tools bieden (woorden zoeken, woorden tellen, eventueel met optie zelf lijsten van woorden die uitgesloten moeten worden te leveren) zouden handig zijn.
4. De onder 3. beschreven functionaliteiten zijn echter beperkt. Het is belangrijk dat er ook inzicht is in waar de termen betrekking op hebben en in welke context ze gebruikt worden. Hierbij moet ook de verschuiving van betekenis van woorden in acht genomen worden. Twee voorbeelden die hierbij naar voren kwamen waren:
 - a. De aanwezigheid van het woord 'held' wil nog niet zeggen dat de persoon waarover de biografie gaat een held genoemd wordt (het kan de held in zijn roman zijn, of zijn tegenstander kan juist een held genoemd worden).
 - b. Het noemen van Amerika of plaatsen in Amerika geeft niet per definitie aan of iemand een band met het land heeft of niet. Het kan vaak genoemd worden in verband met een zakenman die overal ter wereld zaken doet en slechts een maal expliciet over iemand die er op jonge leeftijd naar toe emigreert en er later een belangrijke functie uitoefent.

5. Er moeten filters of andere methoden komen om 'ruis' uit bepaalde queries te kunnen halen.

Andere uitdagingen hebben betrekking op het modelleren van kennis waarbij meerdere aspecten samen komen:

6. Een historicus begint altijd met voorkennis en gebruikt die voor interpretatie. Er is (nogmaals) een grote behoefte tot het op verschillende manieren ordenen van de data en de biografieën.
7. RDF lijkt niet praktisch voor het modelleren van perioden, wel voor het modelleren van context.
8. Informatie over de bronnen moet duidelijk naast de zoekresultaten weergegeven worden. Idealiter kan ook op metadata over de bronnen geselecteerd worden (bijvoorbeeld om beschrijvingen uit verschillende periodes met elkaar te vergelijken).
9. Er moet inzicht zijn in niet alleen welke bronnen de demonstrator gebruikt, maar ook wat hij met de geselecteerde data doet
10. Representaties, waar mogelijk op kaarten, combinaties van plaats- en tijdsaanduidingen, zijn behulpzaam om de bewegingen van mensen in kaart te brengen.
11. Moeilijke vragen zijn wellicht nooit echt te beantwoorden met behulp van de computer, maar de computer kan wel patronen laten zien die richting geven aan een hypothese gerelateerd aan een moeilijke vraag.

Interpretatie van de observaties

De observaties die tijdens de workshop door de deelnemers gemaakt zijn, bieden deels inzicht in wenselijke eigenschappen van de tool die binnen BiographyNet ontwikkeld moet worden en deels in hoe taaltechnologie ondersteuning kan bieden.

Wenselijke eigenschappen en mogelijkheden van de tool

De tool moet meerdere mogelijkheden bieden om informatie te selecteren en weer te geven. Idealiter kunnen gebruikers ook zelf eigenschappen aangeven waarop gezocht kan worden, zodat mogelijkheden niet beperkt blijven tot voorgedefineerde zoekvelden. Verder zijn naast het zoeken door gestructureerde data ook klassieke zoekfuncties handig. Hierbij kan taalanalyse bij het vaststellen welke termen meegenomen moeten worden van dienst zijn en zou het handig zijn als gevonden termen in de tekst door highlights worden weergegeven. Aanwezigheid van geavanceerde versies van "word count" kan ondersteuning bieden bij historiografisch onderzoek. Hierbij is het van belang dat de gebruiker de mogelijkheid heeft zelf woorden aan te geven die genegeerd of juist geteld moeten worden.

Taaltechnologische ondersteuning

Taaltechnologie kan binnen BiographyNet voor twee taken worden ingezet. Ten eerste proberen we met behulp van taaltechnologie informatie uit teksten te halen en die in RDF te representeren. Ten tweede kan taaltechnologie ingezet worden, om te kijken hoe over mensen geschreven wordt.

Uit de observaties van de eerste opdrachten blijkt dat het identificeren van informatie in tekst en dit weer te geven in RDF een nuttige bijdrage levert. Hierbij levert informatie die goed in RDF weergegeven kan worden en waarvoor we technologie in huis hebben zoals beroepen, plaats- en tijdsaanduidingen en eigenschappen die expliciet aan personen worden toegedicht al een bijdrage. Door deze informatie in RDF weer te geven, kan informatie goed doorzocht worden, kunnen individuen op basis van deze informatie voor groepsanalyse geselecteerd worden, etc.

In het tweede deel van de workshop, kwam met name het belang van context en gedetailleerde analyse aan bod. Het onderzoeken van woordgebruik alleen is niet genoeg: er moet gekeken worden waar het woordgebruik betrekking op heeft. Dit betekent dat ook voor historiografisch onderzoek analyse op zinsniveau, waarbij relaties tussen entiteiten en waardeoordelen expliciet worden weergegeven, van belang is. Verder moet er rekening mee gehouden worden dat betekenissen door de tijd veranderen en er niet zomaar aangenomen kan worden dat toenemend of afnemend gebruik van een woord *an sich* iets over historiografie zegt. We kunnen dit probleem aanpakken door technieken die betekenis op basis van distributie weergeven op de teksten in de biografieën en vergelijkend materiaal uit dezelfde periode toe te passen. Hiermee zouden verschuivingen in betekenis in kaart gebracht kunnen worden. Tenslotte ligt er bij historici de uitdaging te ontdekken welke termen representatief zouden kunnen zijn voor een bepaalde visie. Een van de groepen stelde voor hierbij eerst naar teksten waarvan bekend is dat ze deze visie uitdragen (bijvoorbeeld nationalistische teksten) te beginnen en te kijken of er gelijkenissen zijn tussen deze teksten en biografieën. Statistische methoden kunnen hiervoor ingezet worden.

De eerst volgende stappen in de taaltechnologie binnen BiographyNet zullen zich richten op het identificeren van informatie over personen. Hierbij zal de nadruk liggen op het creëren van een tijdlijn van het leven van de individu, waar opleiding, beroepen en woonplaatsen in kaart gebracht worden. Het tweede aandachtspunt is het identificeren van personen en relaties tussen personen in biografieën. Deze informatie kan relatief goed geïdentificeerd worden en eenvoudig weergegeven worden in RDF.

Voor het beantwoorden van historiografische vragen met ondersteuning van taaltechnologie is nauwgezette samenwerking tussen geschiedkundigen en taalkundigen in eerste instantie noodzakelijk. De eerst volgende stap zal hier dan ook zijn om gezamenlijk aan een of een paar use cases te werken en op basis daarvan te kijken welke technieken aan geschiedkundigen aangeboden kunnen worden.

Vervolg van de Workshop

De workshop heeft de onderzoekers van BiographyNet het nodige materiaal gegeven om de volgende stappen binnen het project te kunnen zetten. Het inzicht dat deze workshop heeft opgeleverd is voldoende om een eerste versie van de tool die binnen het project voor geschiedkundigen ontwikkeld zal worden op te leveren. Wij streven ernaar een eerste versie van de tool in de herfst van 2014 af te hebben. Hiermee kan een vervolg op deze workshop plaatsvinden waarbij geschiedkundigen (beter) inzicht kunnen krijgen in de mogelijkheden die door taaltechnologie en Semantic Web tools geboden worden. Deze vervolgworkshop zal daarnaast gericht zijn op het verbeteren van de tool en het onderzoeken van volgende stappen in het project.

Bijlage 1: De Uitnodiging

Workshop Geschiedenis en Digitale Historische Tekstanalyse – 4 April 2014, Vrije Universiteit

De laatste jaren is er veel aandacht voor de mogelijkheden van ‘Digital Humanities’, of ‘e-Humanities’, voor historisch onderzoek. Bij Digital Humanities projecten is er vrijwel altijd sprake van een nauwe samenwerking tussen experts van de geesteswetenschappen en de bètawetenschappen. Een cruciaal onderdeel van een dergelijke interdisciplinaire samenwerking is het leren van elkaars taal en onderzoeksmethoden.

Aan de VU loopt sinds 2012 het Digital Humanities project BiographyNet, dat zich richt op het ontwikkelen van een tool voor het analyseren van de biografieën in het Biografisch Portaal van Nederland (<http://www.biografischportaal.nl>), waarin biografische woordenboeken van de achttiende tot de eenentwintigste eeuw zijn opgenomen. Er kan een rijke variëteit aan historische vragen gesteld worden aan het Portaal, met betrekking tot bijvoorbeeld prosopografisch onderzoek (wat kun je zeggen over de carrières van de gouverneurs-generaal van Nederlands-Indië?), historiografische kwesties (hoe werden gouverneurs-generaal in de loop der tijd beoordeeld?) en thematische onderwerpen (welke personen speelden een belangrijke rol in de koffiehandel?). Voor een vruchtbare analyse met behulp van digital humanities technieken is het echter ook zaak om dergelijke vragen te vertalen naar voor de computer herkenbare elementen. In deze workshop hopen wij op een uitwisseling van ideeën die ons verder helpt bij het ontwikkelen van de BiographyNet tools en die de deelnemers een goed beeld geeft van de mogelijkheden van digitale historische tekstanalyse. Na een introductie worden de deelnemers gevraagd om in groepjes met behulp van het Biografisch Portaal eerst een eenvoudige en vervolgens een moeilijke historische vraag te benaderen. Welke informatie kan en moet het Biografisch Portaal leveren? Wat zijn de beperkingen? Hoe kunnen historische vragen vertaald en opgedeeld worden in bouwstenen die te analyseren zijn met behulp van technieken uit de digital humanities?

Programma:

13:00 – 13:15: Inloop

13:15 – 14:00: Introductie Antske Fokkens (computationale linguïstiek) en Serge ter Braake (geschiedenis)
+ vragen

14:00 – 14:45: Werkgroep sessie 1 (eenvoudige vraag)

14:45 – 15:15: Pauze

15:15 – 16:00: Werkgroep sessie 2 (moeilijke vraag)

16:00 – 16:30: Uitkomsten en discussie

16:30 -17:00: Borrel

Aanmelding tot en met zondag 31 maart via deze [link](#). De exacte locatie wordt nader bekend gemaakt. Mocht je van tevoren graag een bepaalde onderzoeksvraag behandeld zien die (deels) te beantwoorden is met behulp van het Biografisch Portaal laat het dan weten aan een van de organisatoren.

Hopelijk tot 4 april,

Antske Fokkens (a.s.fokkens@vu.nl)

Serge ter Braake (s.ter.braake@vu.nl)

Bijlage 2: De vragen

I A: Mensen en beroepen: de gouverneurs-generaal van Nederlands-Indië

Tussen 1610 en 1949 was er een Nederlandse gouverneur-generaal in Nederlands-Indië, als hoogste vertegenwoordiger van eerst de Republiek en uiteindelijk het Koninkrijk der Nederlanden. Het was zijn taak ervoor te zorgen dat het bestuur en de rechtspraak over de verschillende territoria in Nederlands-Indië goed werden geadmistreerd. In de loop der eeuwen is het profiel van deze gouverneurs veranderd, vaak samenhangend met de politieke omstandigheden in Nederland.

In deze opdracht wordt jullie gevraagd te bedenken hoe we een groepsportret kunnen destilleren van deze groep mensen aan de hand van het Biografisch Portaal. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Hoe achterhalen we op de meest efficiënte manier wie de gouverneurs-generaal van Nederlands-Indië waren?
- Wat willen weten over hun 'externe' karaktertrekken zoals geboorteplaats, leeftijd bij benoeming et cetera? Is die informatie ook uit de tekst te halen?
- Willen we ook nog iets weten over hun 'interne' karaktertrekken?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

I B: Mensen en prestaties: beroemde sporters uit de geschiedenis

De status van sporters is door de tijd heen aan verandering onderhevig geweest. Bij de Grieken waren ze grote volkshelden, in negentiende-eeuwse biografische woordenboeken komen ze niet voor. In de twintigste eeuw kregen sporters weer een nationale heldenstatus. In 2004 wist voetballer Johan Crujff zelfs plaats nummer 6 te verwerven in de verkiezing voor de Grootste Nederlander en kreeg Marco van Basten ook een respectabele 25^e plaats.

In deze opdracht wordt jullie gevraagd te bedenken hoe we een groepsportret kunnen maken van beroemde sporters uit de Nederlandse geschiedenis aan de hand van het Biografisch Portaal.

De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Welke soort sporters worden blijkbaar belangrijk genoeg geacht om opgenomen te worden in biografische woordenboeken?
- Welke ontwikkeling door de tijd valt waar te nemen?
- Is er iets te zeggen over de verhouding mannelijke/vrouwelijke sporters?
- Kunnen we iets zeggen over de belangrijke sportevenementen (Olympische Spelen; WK's) waar ze bij betrokken waren?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

NIET GEDAAN

I C: Mensen en plaatsen: beroemdheden uit Weert

Weert is een stad in Limburg, grenzend aan Noord-Brabant en België. Om die reden wordt Weert wel de 'poort van Limburg' genoemd. Als grensgemeente is Weert een interessante casus om te kijken naar geografische mobiliteit, en welke rol een dergelijke locatie kan spelen in de grotere Nederlandse geschiedenis. Ook is het interessant om te kijken welke mensen uit Weert zijn opgenomen in de biografische woordenboeken.

In deze opdracht wordt jullie gevraagd te bedenken hoe we informatie over de rol van Weert in de geschiedenis uit het Biografisch Portaal kunnen halen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Hoeveel mensen uit het Biografisch Portaal zijn in Weert geboren/gestorven?
- Kan er iets gezegd worden over de gedeelde *claim to fame* van deze mensen? Hoe zou deze zich verhouden tot de *claim to fame* van andere mensen in het portaal?
- Weert is een grensgemeente, waren de elites daar ook meer 'internationaal'?
- Wordt Weert ook genoemd in andere contexten behalve van die als geboorte- of sterfplaats van een individu?
- Stond Weert tijdens een bepaalde periode meer 'op de kaart' dan een andere?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

I D: Mensen en hun omgeving: Hans van Mierlo

Hans van Mierlo was een van de voornaamste oprichters van D66 en jarenlang het boegbeeld van deze partij. Hij mag gelden als een van de succesvolste politici van de twintigste eeuw. Als men een biografie van Van Mierlo zou willen schrijven zou je je kunnen afvragen welke omstandigheden mogelijkwijs tot zijn succes hebben bijgedragen. Wat kunnen we zeggen over zijn opleiding, sociale afkomst en netwerken?

In deze opdracht wordt jullie gevraagd te bedenken hoe we informatie over Hans van Mierlo en zijn contact uit het Biografisch Portaal kunnen halen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Met welke randvoorwaarden (sociale achtergrond, connecties etc...) werd Hans van Mierlo een succesvol politicus?
- Hoe verhouden zijn meetbare, externe, karaktertrekken zich tot die van andere succesvolle en minder succesvolle politici uit de twintigste eeuw?
- Komt van Mierlo voor in de biografieën van andere personen en zo ja, wat kan er over hen gezegd worden?
- Bij welke belangrijke gebeurtenissen, die ook in andere biografieën genoemd worden, was Van Mierlo betrokken?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

I E: Mensen uit alle tijden: Van Assendelft, De Witt en Nicolai

Het is algemeen bekend dat kinderen van politici vaak in de voetsporen van hun ouders (tot voor kort altijd vaders) treden. De vraag is in hoeverre deze continuïteit verder strekt dan een aantal generaties. Stammen hedendaagse politici rechtstreeks af van regenten uit de Middeleeuwen? De families Van Assendelft (vanaf de vijftiende eeuw), De Witt (vanaf vooral de zeventiende eeuw) en de familie Nicolai (vanaf de zestiende eeuw) waren royaal vertegenwoordigd in de populatie van Nederlandse volksvertegenwoordigers. Is dat nu nog steeds zo?

In deze opdracht wordt jullie gevraagd te bedenken hoe we relevante informatie over de genoemde families op de meest efficiënte wijze uit het Biografisch Portaal kunnen halen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Welke telgen van deze families staan vermeld in het Biografisch Portaal en wat was hun *claim to fame*?
- Hoe kunnen we het beste een stamboom uit de beschikbare gegevens destilleren?
- Hoe voorkom je dat een andere familie met dezelfde naam wordt meegenomen in een zoekvraag?
- Hoe houd je verschillende mensen met dezelfde naam uit elkaar?
- Is het te traceren als een tak doodloopt, of is er dan simpelweg geen biografie beschikbaar?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

NIET GEDAAN

2 A: Mensen en gebeurtenissen: welke personen speelden een sleutelrol in het rampjaar 1672?

1672 is bekend als het Rampjaar voor de Republiek. De Republiek werd door vier mogendheden tegelijk aangevallen, de zittende regenten stonden zwaar onder druk en er werd geschreeuwd om de benoeming van Willem III van Oranje tot stadhouder. Deze omstandigheden leidden er toe dat de positie van Johan de Witt als raadpensionaris onhoudbaar werd en dat hij moest aftreden. Kort daarna werden hij en zijn broer Cornelis vermoord door Haagse schutters en door een woedende menigte aan stukken gescheurd.

In deze opdracht wordt jullie gevraagd te bedenken hoe we personen uit het Biografisch Portaal het beste kunnen koppelen aan de gebeurtenissen (*events*) in 1672. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Welke *events* uit 1672 kunnen we überhaupt destilleren uit het Biografisch Portaal?
- Welke personen worden veelvuldig genoemd in verband met die *events*? (in hun eigen lemma of in die van een ander)
- Is het mogelijk een netwerk van staatsgezinden of oranjegezinden uit het portaal te halen?
- In hoeverre worden events op een andere manier beschreven?
- Geeft het portaal *leads* om informatie elders te zoeken?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

2 B: Mensen en mensbeschrijvingen: Hoe is het beeld van de gouverneurs-generaal van Nederlands-Indië in de loop der eeuwen veranderd?

Het beeld dat we hebben van mensen uit de geschiedenis verandert met de tijd. Veel van de gouverneurs-generaal van Nederlands-Indië hebben met de ogen van nu onmenselijke misdaden begaan. Gouverneur Jan Pietersz Coen bijvoorbeeld, wordt in de negentiende eeuw nog gezien als een strenge, deugdzame, sober levende man en een harde werker, terwijl hij in de twintigste eeuw wordt beschreven als een compromisloze man die zich alleen kon verlaten op bruto geweld. De vraag is in hoeverre dergelijke sentimentveranderingen door de tijd heen zijn te traceren.

In deze opdracht wordt jullie gevraagd te bedenken hoe we de beoordeling van de gouverneurs-generaal van Nederlands-Indië door de tijd heen in kaart kunnen brengen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Wat betekenen termen als 'held' en 'moedig' in de negentiende-eeuwse context?
- Wat voor gebeurtenissen werden eerst niet (of nauwelijks) en later wel opgenomen in iemands biografie?
- Zijn er binnen bepaalde biografische woordenboeken duidelijke verschillen aan te wijzen in de manier waarop auteurs personen uit de geschiedenis beoordelen?
- Wat zegt dit alles voor het gebruik van het Biografisch Portaal als bron?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

2 C: Mensen en naties: In hoeverre was er sprake van 'nationalisme' in biografische woordenboeken van de negentiende tot de eenentwintigste eeuw?

De negentiende eeuw was niet alleen de eeuw van het nationalisme, maar ook die van het biografische woordenboek. Vaak wordt de bloei van de woordenboeken direct gekoppeld aan het bloeiende nationalisme. De vraag is in hoeverre we dan ook 'nationalistische' sentimenten kunnen terugvinden in de biografische woordenboeken.

In deze opdracht wordt jullie gevraagd te bedenken hoe we het voorkomen van nationalisme in biografische woordenboeken in kaart kunnen brengen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Welke termen kunnen als nationalistisch bestempeld worden?
- Hoe kom je aan een goede lijst met termen en uitdrukkingen die iets van Nederlandse nationalisme suggereren?
- Is nationalisme überhaupt te vinden op basis van termen?
- Zouden we zelfs iets van verzuiling kunnen terugvinden in de woordenboeken?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

2 D: Mensen en wereldgeschiedenis: Vanaf wanneer werd de ontdekking van Amerika van werkelijk belang voor de Nederlandse elites?

In 1492 'ontdekte' Columbus per ongeluk Amerika. Een periode van kolonisatie en uitbuiting/uitroeiing van de oorspronkelijke bewoners van dit continent begon, gevolgd door het vervoeren van slaven uit Afrika daar naar toe. Bronnen uit het huidige Nederland laten echter geen verschil zien tussen 1492 en 1493. De belangrijkste Europese ontdekking ging vooralsnog grotendeels langs de Nederlandse elites heen. Pas veel later raakten Nederlandse elites werkelijk betrokken bij de 'Nieuwe Wereld'. De vraag is nu of die ontwikkeling te traceren valt in de biografische data van het portaal.

In deze opdracht wordt jullie gevraagd te bedenken hoe we 'de ontdekking van Amerika' in Nederland in kaart kunnen brengen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Hoe kan je het meest efficiënt zoeken naar vermeldingen van Amerika?
- Wanneer is een melding werkelijk relevant voor een vraagstelling naar de betrokkenheid bij de Nieuwe Wereld?
- Vanaf wanneer hebben mensen hun opname in een biografisch woordenboek te danken aan hun betrokkenheid bij Amerika?
- Welke delen van Amerika worden in de loop der tijd belangrijk of minder belangrijk voor Nederland?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?

2 E: Mensen en mobiliteit: welke patronen kunnen er ontdekt worden in de manier waarop mensen zich geografisch voortbewogen?

Volgens een oud Italiaans gezegde moet je koeien en vrouwen alleen nemen als ze uit je eigen dorp komen. Ook in Nederland waren mensen lange tijd weinig mobiel. Nederlandse elites daarentegen waren minder honkvast. Hoge adel trouwde over de 'landsgrenzen' heen en juristen vestigden zich in het bestuurlijke centrum Den Haag of in een andere stad. Het is interessant om te kijken of de mate van mobiliteit van Nederlanders vast te leggen is door de eeuwen heen.

In deze opdracht wordt jullie gevraagd te bedenken hoe we de geografische mobiliteit van Nederlanders in kaart kunnen brengen. De uitwerking staat vrij, maar denk bijvoorbeeld aan:

- Welke 'buitenlanders' staan er in de biografische woordenboeken?
- Welke 'Nederlanders' die vooral elders resideerden staan in de biografische woordenboeken?
- Hoe kunnen we plaatsnamen die genoemd worden in een biografische tekst op verantwoorde manier koppelen aan iemands geografische mobiliteit?
- Is er een zekere 'beweging' (bijvoorbeeld van Oost naar West) te traceren en in hoeverre is dat wellicht te relateren aan een *bias* in de woordenboeken?
- Hoe kan de benodigde informatie voor deze vragen in RDF gerepresenteerd worden?