

Revolutionaire tools op DH Benelux

# Complexe tekst? Eitje voor TAG!

Het Research & Development-team van het KNAW Humanities Cluster ontwikkelt digitale tools voor geesteswetenschappelijk onderzoek en presenteert dit op de DH Benelux van 6-8 juni.

Thijs van der Veen

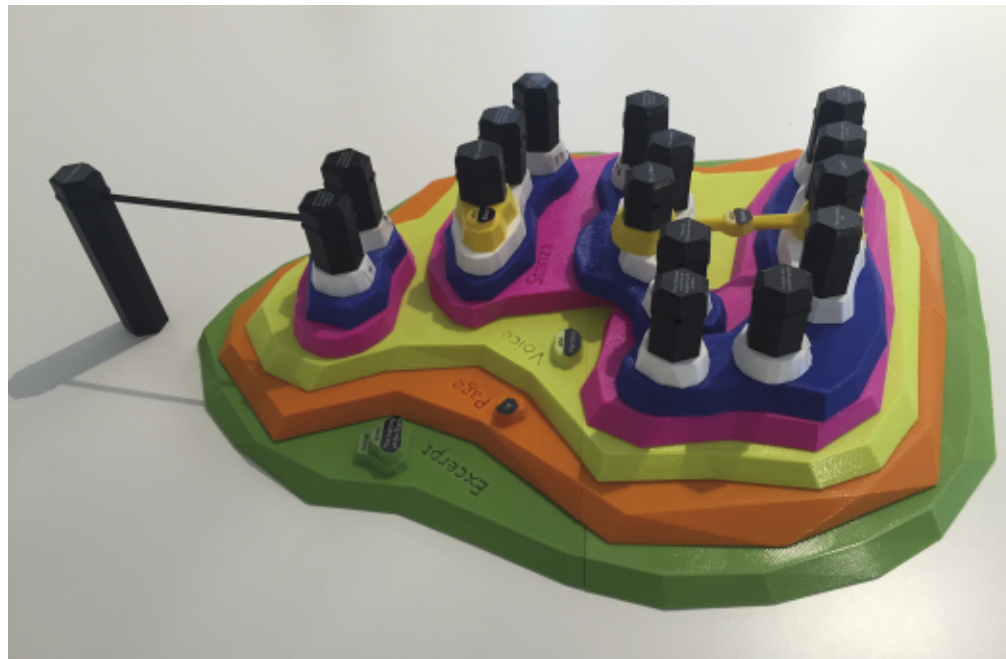
Tekst kan worden gezien als een netwerk van informatie over inhoud, tekstvariatie en structuur. Het is lastig, zo niet onmogelijk, om al deze elementen (en de relaties ertussen) in kaart te brengen met behulp van bestaande tools zoals TEI-XML. Het R&D-team van het Humanities Cluster heeft hiervoor een nieuw model ontwikkeld, Text as Graph (TAG), dat gebruik maakt van een hypergraafstructuur. Teamleider en software ingenieur Ronald Haentjens Dekker: "Met TAG kunnen zeer complexe teksten gemakkelijk worden opgeslagen, bewerkt en bevestigd. TAG implementeert de tekstcollatielool HyperCollate, waarmee teksten met elkaar kunnen worden vergeleken op zeer gedetailleerd niveau (zowel variatie binnen één tekst als variatie tussen verschillende tekstversies). Deze manier van onderzoeken resulteert in een zeer rijke output.

## DH Benelux 2018

De 5e Digital Humanities Benelux Conference vindt plaats van 6-8 juni 2018 op het Internationaal Instituut voor Sociale Geschiedenis (IISG) in Amsterdam.

Tijdens de conferentie komen interdisciplinaire DH-onderzoekers uit België, Nederland en Luxemburg samen om hun onderzoeksresultaten, tools en projecten met elkaar te delen.

Het KNAW Humanities Cluster is dit jaar de organisator.



Fysiek 3D-model van Text as Graph (TAG). Dit specifieke model van TAG laat zien hoe stukken tekst uit *The Hunting Of The Snark* van Lewis Carroll tegelijkertijd onderdeel uitmaken van elementen als bijvoorbeeld een excerpt, een page, een voice en een stanza.

Credits R&D-team KNAW Humanities Cluster

Het menselijk brein kan daardoor overweldigd worden, visualisatie van de output kan uitkomst bieden. Tijdens DH Benelux gaat het R&D-team in op de uitdagingen die dat met zich meebrengt in de lezing TMI? Visualisation as Research Instrument for Computational Philology.

## Digitaal onderzoek

Een tweede lezing van het R&D-team tijdens de DH Benelux-conferentie is qua onderwerp minder technisch maar voor de toekomst van digitaal onderzoek minstens zo relevant. Marijn Koolen (R&D), Jasmijn van Gorp (Universiteit Utrecht) en Jacco van Ossenbruggen (CWI & VU) bespreken naar aanleiding van hun paper 'Lessons Learned from a Digital Tool Criticism Workshop' hoe wetenschappers in hun onderzoek het gebruik van digitale tools kunnen verbeteren.

Reflectie is het toverwoord en dan het liefst in een groep. Nu is samenwerking in digitaal geesteswetenschappelijk onderzoek sowieso aan te raden, omdat daar immers veel verschillende vaardigheden bij komen kijken. Traditionele geesteswetenschappers opereren meestal alleen, maar door in een groep te reflecteren op de denkstapen en keuzes die eenieder maakt bij het gebruik van digitale tools, kom je tot nieuwe inzichten en beter onderbouwde afwegingen in hoe, wanneer en waarom bij het inzetten van tools. Wellicht voor de hand liggende vragen, maar in de praktijk blijkt dat onderzoekers dit bewust of onbewust nog te weinig doen. Daarom organiseren Jacco, Jasmijn en Marijn tijdens DH Benelux 2018 hier een workshop over.

<http://2018.dhbenelux.org>

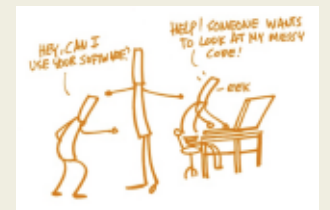
## GELEZEN

*The Open Science Training Handbook*

Marjan Grootveld

Het "Open Science Training Handbook" is het resultaat van een week samen schrijven – en van de ervaring van veertien auteurs, die betrokken zijn bij ondermeer FOSTER, CESSDA en 4TU.ResearchData. Het handboek is bestemd voor trainers en frist in hoofdstuk twee de kern op van learning, teaching en training. Hoofdstuk één legt de inhoudelijke basis met informatie over onder andere open data, open software, citizen science en reproducible research. Het beschrijft systematisch de voornaamste kennis en vaardigheden, leerdoelen en eventuele obstakels en misvattingen. De laatste hoofdstukken zijn uitgesproken praktisch: na een checklist voor het organiseren van trainingsbijeenkomsten volgt een hoofdstuk met ruim twintig opdrachten. Hieronder zijn zowel ijsbrekers van enkele minuten als discussieopdrachten van een half uur, allemaal toegesneden op Open Science en met tips voor de uitvoering, waaronder "This exercise should be fun". Bij toepassing in een OpenAIRE-workshop bleek dit helemaal te kloppen. Handboek en literatuurlijst zijn online beschikbaar.

[book.fosteropenscience.eu](http://book.fosteropenscience.eu)



## COLUMN

### Een mooie taak voor het netwerk van bibliotheken

Het gaat goed met de kranten. De lezersaantallen stijgen, vooral online. Die ontwikkeling is voor een belangrijk deel te danken aan nepnieuws. Door alle commotie rond nepnieuws en de schandalen rond Cambridge Analytica en Facebook zijn steeds meer mensen zich ervan bewust hoezeer zij gemanipuleerd worden. En dus keren zij terug naar oude, vertrouwde nieuwsbronnen: dagbladen en de website van bijvoorbeeld de NOS.

Voor mij geldt hetzelfde: ik vertrouw minder op de nieuwsfeeds van Google en begin en eindig mijn dag meestal op de NOS-site. Met Facebook heb ik nooit veel gedaan en vanwege Cambridge Analytica heb ik mijn account verwijderd – je moet ergens een grens trekken.

We zien dus een teruggaande beweging naar bronnen die hun betrouwbaarheid in het verleden hebben bewezen.

Ik denk dat we een vergelijkbare ontwikkeling zullen zien bij grote bibliotheken. Lang golden die als de kennisbewaarplaatsen bij uit-

stek. Een goede bibliotheek beschikte over de nieuwste handboeken, de beste en meest relevante tijdschriften en de bibliothecaris of balie medewerker (m/v) kon desgevraagd optreden als gids.

De afgelopen vijftien jaar zijn bibliotheken op grote schaal gaan digitaliseren, al dan niet in samenwerking met Google. Ik ga zelf alleen nog naar een bibliotheek als ik boeken moet inzien die nog niet online digitaal beschikbaar zijn.

Toch verwacht ik dat de komende jaren meer mensen lid zullen worden van een bibliotheek. Althans: als die zich, net als kranten, beter zullen positioneren als bronnen

van betrouwbare kennis. Wat ik nu zie is dit: een bibliotheek als de Koninklijke Bibliotheek (KB) zet zich onder meer in om grote datasets beter te kunnen analyseren met nieuwe technologieën en kunstmatige intelligentie. Dat is lovenswaardig, maar volgens mij zou het tegelijkertijd nuttig zijn om 'gewone' gebruikers – scholieren, studenten,



foto Leo van Velzen

ouderen – te leren hoe zij bijvoorbeeld nepnieuws kunnen onderscheiden van echt nieuws. En hoe je kunt vaststellen of informatie op een site betrouwbaar is of niet. Gewoon slimmer en creatiever zoeken op internet dus.

Dat is natuurlijk geen taak van de KB alleen. Er bestaat een netwerk van openbare bibliotheken. De hele infrastructuur ligt er dus om bijvoorbeeld op lokaal niveau workshops aan te bieden hoe je optimaal gebruik kunt maken van internet en van gedigitaliseerde bronnen van onder meer bibliotheken. Je zou hier ook filmpjes over kunnen maken (er zijn er al een paar), maar in mijn ervaring werken workshops beter, want die zijn interactiever. Overigens zouden dergelijke workshops volgens mij niet alleen zinvol zijn voor scholieren, studenten en ouderen. Ik kom ook geregeld docenten en zelfs bibliothecarissen tegen die baat zouden kunnen hebben bij zo'n workshop of bijscholing.

Internet is een uitdijend hee-

waar je de prachtigste maar ook de onzinnigste informatie kunt vinden. Om niet te worden misleid door nepnieuws, keren mensen terug naar kranten en naar nieuwssites die zij vertrouwen. "De bibliotheek maakt je slimmer, vaardiger en creatiever", luidt de missie van sommige bibliotheken, waaronder de KB. Ik vind dat een mooie missie, die zou kunnen worden verwezenlijkt als het netwerk van bibliotheken cursussen zou gaan aanbieden over hoe wij optimaal gebruik kunnen maken van internet. Voor jong en oud, voor beginners en gevorderden.

Ik vermoed dat daar meer belangstelling voor zal bestaan dan voor slimme tools met de nieuwste technologieën, hoewel ook die moeten blijven worden ontwikkeld.

Ewoud Sanders

Taalhistoricus en journalist.

Sanders is vaste medewerker van onder meer NRC Handelsblad en Onze Taal.