

ISEBEL

Op 1 september 2017 is er een nieuw project bij het Meertens Instituut gestart genaamd ISEBEL¹, wat staat voor: Intelligent Search Engine for Belief Legends. Nu er her en der meer databases ontstaan met volksverhalen en folklore wordt het immers interessant om data te gaan combineren. Het project is dan ook een zogenaamd Trans-Atlantic Digging Into Data project² waaraan vooralsnog twee andere partners deelnemen. Er moet een ‘harvester’ gebouwd worden waarmee simultaan gegevens kunnen worden geoogst uit de Nederlandse Volksverhalenbank³, de Deense Volksverhalenbank⁴ en de Noord-Oost Duitse folklorebank (die luistert naar de naam WossiDia⁵). Het project wordt geleid door Theo Meder in samenwerking met de afdeling Technische Ontwikkeling van het Meertens Instituut, terwijl de buitenlandse partners zijn: Tim Tangherlini van de Universiteit van California (UCLA) en Christoph Schmitt en Holger Meyer van de Universiteit van Rostock (Mecklenburg). De nationale en regionale databases blijven gewoon intact, maar stellen hun gegevens tevens beschikbaar aan de *search engine* van ISEBEL die nu ontwikkeld wordt.

Het ophalen van de diverse gegevens is al geen sinecure: de zoekmachine moet om te beginnen weten in welke velden welke informatie te vinden is. Als eerste pilot wordt nu gewerkt aan het ophalen van de volgende gegevens uit de drie databases: het verhaal, de verteller, de plaats van vertellen, de datum van vertellen, en de *keywords*. Om de harvester zo algemeen bruikbaar mogelijk te maken, worden de resultaten in velden gezet waar internationaal consensus over is – de velden zullen met andere woorden zo veel mogelijk voldoen aan de Dublin Core standaard.⁶ Dit is belangrijk met het oog op verdere uitbreiding met meer databases: er is reeds belangstelling getoond vanuit IJsland, Noorwegen en Zweden. Het is de bedoeling dat ISEBEL wordt voorzien van geografische visualisaties, zodat op kaarten de verspreiding van bepaalde fenomenen goed te zien is. Alleen al met Nederland, Denemarken en Noord-Oost Duitsland zal dit interessante verspreidingskaarten opleveren, maar bij uitbreiding met meer Scandinavische landen wordt het alleen maar boeiender. En natuurlijk zouden we graag in een later stadium ook het materiaal uit de Vlaamse Volksverhalenbank, de Catalaanse en de Portugese databases in het project willen betrekken, maar dat pas in een stadium dat ISEBEL naar behoren functioneert.

Bij een internationale harvester gaan we natuurlijk op taalproblemen stuiten. Ook als ISEBEL weet *waar* de gegevens moeten worden opgehaald, dan nog moet ook duidelijk zijn *welke* gegevens er precies opgehaald moeten worden. De verschillende trefwoordsystemen moeten aan elkaar gekoppeld worden, zodanig dat ISEBEL weet dat bij een Engelstalige zoekvraag naar “death” er in het Nederlands gezocht moet worden naar “dood”, in het Fries naar “dea”, in het Duits naar “Tod”, in het Deens naar “død”, in

1 De website van ISEBEL, www.isebel.eu, bevat nu alleen nog wat basale informatie over het project, maar de site zal op zeker moment de plek voor de zoekmachine worden.

2 Alle door NWO toegekende Trans-Atlantic Digging into Data projecten worden hier opgesomd: <https://www.nwo.nl/actueel/nieuws/2017/gw/zestien-internationale-onderzoeksfinanciers-kondigen-de-winnaars-van-de-2017-t-ap-digging-into-data-challenge-aan.html>

3 Zie www.verhalenbank.nl

4 Hier te vinden: <http://etkspace.scandinavian.ucla.edu/etkSpace/index.php> en <http://etkspace.scandinavian.ucla.edu/danishfolklore/#>

5 Zie <https://apps.wossidia.de/webapp/run>

6 Hierover https://nl.wikipedia.org/wiki/Dublin_Core

het Zweeds naar “död” en in het IJslands naar “dauða”. In dit geval zou men dit nog aan Google Translate kunnen overlaten, maar bij meer specialistische volkskundige begrippen als kabouter, alf, plaaggeest en spookpaard gaat dit al snel problemen opleveren. Er zal dus een uitgebreide thesaurus voor meerdere talen opgesteld moeten worden – dat kan deels automatisch, maar moet ook deels met de hand. Als proef zal het team de databases als eerste gaan bevragen op het begrip “heks”. Over een grote kuststrook van Noordzee en Oostzee zullen – als het allemaal lukt – verspreidingskaarten getekend kunnen worden over het fenomeen “heks”. Wat voor soort verhalen worden er over heksen verteld? Komt het verhaal van de heks die het karnen onmogelijk maakt overal even vaak voor? Moeten heksen overal regelmatig iemand doden? Is het fenomeen van de kattendans internationaal? Hoe evident is het pact met de duivel? Als we informatie over de vertellers uitbreiden met het geslacht van de vertellers, tekent er zich dan een verschil af tussen verhalen van mannen en verhalen van vrouwen? Indien we aan de plaats van vertellen de plaats van handelen toevoegen, is dan te zien hoe dicht verhaalde gebeurtenissen zich bij de eigen leefomgeving afspeelden? ISEBEL zal zich in elk geval vooral concentreren op traditionele sagen, dus voornamelijk op het geloof in bovennatuurlijke fenomenen in het verleden: meerminnen, zeemonsters, de nachtmerrie, magie, waarzeggers, duivelbanners, weerwolven, trollen, tovenaars en wat dies meer zij. ISEBEL heeft de potentie en de ambitie om uiteindelijk uit te groeien tot een Europese zoekmachine op het gebied van sagen (en legenden).

Een complicatie die zich nog zal gaan aftekenen, is de periode waarin het verhaalmateriaal verzameld is. De collecties stammen – zeker als we de hoeveelheid databases gaan uitbreiden – lang niet allemaal uit dezelfde tijd. De Deense database is gebaseerd op het materiaal dat is verzameld door Evald Tang Kristensen (1834-1929), terwijl de Duitse data afkomstig zijn uit de verzamelarbeid van Richard Wossidlo (1859-1939). We hebben het in deze gevallen dus over volksverhalen uit het einde van de 19e en het begin van de 20e eeuw. Uit die periode heeft Nederland ook wel materiaal, met name dankzij verzamelaars als G.J. Boekenoogen en C. Bakker, maar de honderden verhalen vallen een beetje in het niet vergeleken met de duizelingwekkende hoeveelheden gegevens die Kristensen en Wossidlo verzamelden. Wat dat betreft kunnen de Nederlanders beter uitwijken naar de vele duizenden verhalen die tijdens het Amsterdamse verzameloffensief onder leiding van J.J. Voskuil zijn bijeengebracht, onder andere door de Fries A.A. Jaarsma. Maar dit speelde in de jaren zestig en zeventig van de 20e eeuw. Globaal gesproken werden bij de Denen en de Duitsers nog bovennatuurlijke verhalen opgetekend waarin mensen geloofden, terwijl de meeste Nederlandse vertellers in hun verhalen terugkijken op fenomenen waar men toen al nauwelijks meer in geloofde. En dit probleem zal zich bij uitbreiding van ISEBEL blijven voordoen: niet elke verhalenbank in Europa bevat evenveel materiaal uit dezelfde periode. Bij de presentatie van de zoekresultaten in ISEBEL zal men zeker rekening moeten houden met de ouderdom van het gevonden materiaal.

Theo Meder