

Kwantitatief onderzoek naar orale cultuur

Zoekmachine herkent allerlei volksdeuntjes

Vroeger stond op menig schoorsteenmantel een bellenspeelklok. Die gaf niet alleen de tijd aan, maar speelde ook elk kwartier een deuntje. Een zoekmachine helpt deze melodieën herkennen.

Erica Renckens

Marieke Lefeber is onderzoekster aan het Meertens Instituut en conservator bij Museum Speelklok in Utrecht. Aan de hand van de bellenspeelklokken onderzoekt zij de muzikale smaak van de elite in de 18^e eeuw. Om de melodieën op de bellenspeelklokken te identificeren, maakte zij gebruik van de Nederlandse Liederenbank van het Meertens Instituut, sinds 2007 online. “De Nederlandse Liederenbank bevat metadata van ruim 150.000 liederen uit de middeleeuwen tot nu”, vertelt Martine de Bruin, projectleider Etnologie aan het Meertens Instituut.

Gelijkenissen en varianten

Ze werkte mee aan het CATCH-project Witchcraft waarin een melodieënzoekmachine voor de Liederenbank ontwikkeld is. “In Witchcraft hebben we gezocht naar een computationele methode om gelijkenis tussen melodieën vast te stellen. We wilden niet alleen exacte matches vinden, maar ook varianten van zo’n melodie.” In het project zijn eigenschappen die volksliedkundigen gebruiken om melodieën te herkennen, omgezet naar algoritmes waarmee de afstand tussen melodieën van volksliederen kan worden berekend.



Martine de Bruin (l) en Marieke Lefeber in Museum Speelklok te Utrecht foto Bart van Vliet

“Met behulp van Witchcraft heeft Marieke Lefeber diverse klinkende melodietjes kunnen identificeren”, aldus De Bruin.

In de Nederlandse Liederenbank zijn naast opnames ook muziekboekjes opgenomen, met daarin melodieën waarvan geen opnames

bestaan. Met behulp van crowdsourcing worden deze muziekboekjes getranscribeerd. De Bruin: “In onze nieuwe muziekeditor kan iedereen de muziek bij een gedeelte van de liedjes uit de Liederenbank bekijken, in transcriptie of in scan.”

VERVOLG OP PAGINA 2

Een veilig huis voor je data

Voor consumenten bestaan keurmerken om bijvoorbeeld gezonde producten te herkennen. Voor onderzoekers is het Data Seal of Approval steeds vaker het keurmerk voor veilig opgeslagen data.

Marika de Bruijne

Data kun je op verschillende manieren opslaan. Op je computer, op de universiteitserver, of in een digitaal nationaal archief. Sommige onderzoeken worden zelfs op YouTube gezet. De eerste optie, je eigen computer, is wel zo makkelijk. Maar veilig is het niet. Hoe weet je als onderzoeker waar je data wel goed kunt opslaan en hoe kun je dat beoordelen?

Het Data Seal of Approval (DSA), oorspronkelijk een initiatief van DANS, is een keurmerk dat verklaart dat de data-archieven die het keurmerk dragen, data veilig opslaan. Het keurmerk is momenteel aan acht data-archieven toegekend. In Nederland zijn dit EASY van DANS

en het Taal Archief van het Max Planck Instituut. De eisen van het keurmerk bepalen onder andere dat er speciale maatregelen moeten worden genomen om de veiligheid van de data, in de brede zin, te waarborgen. De dataopslag moet voor lange termijn geregeld zijn en procedures rondom gebruik van de data dienen transparant te zijn voor de buitenwereld.

Juist en integer

Ook is er veel aandacht voor het beschikbaar stellen van data aan andere academici. “Het DSA geeft aan dat onderzoekers erop kunnen vertrouwen dat de gegevens die ze ontvangen, grondig zijn beoordeeld op juistheid en integriteit”, vertelt Mary Vardigan, één van de board members van DSA en Assistent Directeur van ICPSR, het grootste digitaal archief in de sociale wetenschappen in de Verenigde Staten. “Er is aandacht geschonken aan de bruikbaarheid van de gegevens en het beschikbaar maken van meta-

data die nodig zijn voor het maken van analyses.”

Bekendheid groeit

Momenteel lopen zo’n dertien aanvraagprocedures en het is te verwachten dat het aantal uitgegeven keurmerken zich in 2013 minimaal zal verdubbelen. Dit betekent dat onderzoekers wereldwijd het DSA-logo steeds meer zullen tegenkomen. CentERdata is één van de organisaties die het DSA-keurmerk recentelijk heeft aangevraagd voor de data van het LISS panel. Deze data zijn zowel via een eigen website als via DANS toegankelijk voor alle academici. “Voor onze gebruikers is het belangrijk dat ze er van op aan kunnen dat onze data werkelijk betrouwbaar zijn, het keurmerk geeft dat op een simpele manier aan. Voor onszelf is het keurmerk een erkenning van alle moeite en tijd die we steken in de kwaliteit van onze data en documentatie”, vertelt Annette Scherpenzeel, projectmanager van het LISS panel.



Het DSA-keurmerk

Eind 2012 werd de eerste internationale conferentie over Data Seal of Approval gehouden. Er blijkt veel behoefte te zijn aan een universele en toegankelijke procedure rondom datacertificatie. Het doel voor de nabije toekomst is dan ook om de bekendheid van het DSA te vergroten onder zowel de archieven zelf als onder de dataproducenten en -gebruikers. “Ik hoop dat het keurmerk zich internationaal snel zal verspreiden, want het is een geweldige manier om op een transparante wijze te tonen dat je voldoet aan de normen van digitale duurzaamheid”, vindt Vardigan.

www.datasealofapproval.org



E-DATA & RESEARCH

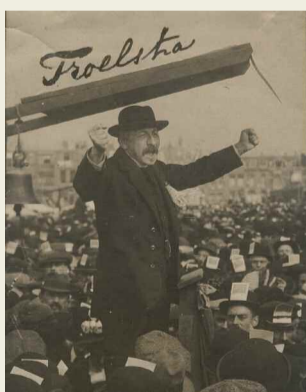
Jaargang 7 | nummer 2

Nieuwsbrief over data en onderzoek in de alfa- en gamma-wetenschappen.

E-data & Research verschijnt drie keer per jaar en wordt mogelijk gemaakt door: CBS, CentERdata, CLARIN-NL, DANS, Huygens ING en de Koninklijke Bibliotheek.

INHOUD

- 2 Interessante bijeenkomsten in de Agenda
- 3 Nederlands vrijgedrag in kaart gebracht
- 3 Krantencorpus online toegankelijk gemaakt
- 4 Deze datasets zijn sinds kort beschikbaar
- 5 De winnaars van de Dataprijs aan het woord
- 6 Wat is waar in onze digitale werkelijkheid?
- 7 Klopt het verhaal bij de beelden?



8 Geen geheimen meer voor wetenschappers

8 Ewoud Sanders is verslaafd aan e-data



Scan deze QR code met een smartphone om de website van E-data te bezoeken.
www.edata.nl