

Online toegankelijk maken en delen van data bekroond en beloond

Liedererenbank en DINED winnaars Dataprijs 2014

De Nederlandse Dataprijs voor het toegankelijk maken en delen van onderzoeksdata gaat dit jaar naar de Nederlandse Liedererenbank en DINED. *Heidi Berkhout*

Op 24 oktober jl. werd voor de derde keer de Nederlandse Dataprijs uitgereikt. € 7.500, een beeld en een diner voor de onderzoeksgroep, met die prijzenpot gingen de Nederlandse Liedererenbank (winnaar Dataprijs humaniora en sociale wetenschappen) en DINED (winnaar Dataprijs exacte en technische wetenschappen) naar huis.

Nog lang niet klaar

“We zijn heel erg blij met de Dataprijs!”, vertelt Martine de Bruin, lid van het projectteam van de Nederlandse Liedererenbank. “De jury noemt onze database een ‘grote aanwinst voor zowel het Nederlands academisch als cultureel erfgoed’. Dat is een bevestiging dat we op het goede spoor zitten. Het prijzengeld besteden we aan onze melodieënzoekmachine. Hiermee zijn beschreven melodieën te vergelijken en kunnen titelloze of gevarieerde melodieën worden geïdentificeerd. En daarvoor is veel data nodig. Het prijzengeld besteden aan een pilotproject lijkt ons een goede voorbereiding op een grotere subsidieaanvraag. Maar misschien wordt het ook nog iets anders. Er zijn nog zoveel verrijkingen en verdiepingen mogelijk: de Liedererenbank is nog lang niet klaar!” Juryvoorzitter prof. Kees Aarts: “In de praktijk wordt de databank al door veel wetenschappers binnen de digital humanities gebruikt voor literair-, muziek-, godsdienst-, algemeen historisch en etnologisch onderzoek. Ook journalisten, musici en middelbare scholieren gebruiken de databank veel. Deze goed gestructureerde databank is daarom een grote aanwinst voor zowel het Nederlands academisch als cultureel erfgoed.”

Begrijpelijke data

Ook Johan Molenbroek en Marijke Dekker, de mensen achter DINED, zijn verrast en blij: “Ons doel is altijd geweest om onbegrijpelijke data



Prof. Louis Grijp (Meertens Instituut) ontving de Dataprijs voor de Nederlandse Liedererenbank, een database van 170.000 Nederlandse liederen van de middeleeuwen tot de 20e eeuw. Ir. Marijke Dekker is namens dr.ir. Johan Molenbroek (TU Delft) voor DINED aanwezig, een database met meetgegevens over de variatie van menselijke vorm en afmetingen foto Annemiek van der Kuil

begrijpelijk te maken. En dat lukt: met onze interactieve tool wordt de variatie in de afmetingen van het menselijk lichaam duidelijk, zonder dat de gebruiker veel kennis over anatomie en statistiek nodig heeft. Naast studenten gebruiken ook bedrijven en ziekenhuizen onze database. De vele vragen geven steeds weer aanleiding om verbeteringen aan te brengen en nieuwe data toe te voegen, zoals menselijke krachten. Onze faculteit Industrieel Ontwerpen van de TUDelft stelt ons steeds weer in de gelegenheid om het benodigde programmeerwerk door een student-assistent te laten uitvoeren. En door het winnen van de Dataprijs kunnen we nu ook een paar grotere, al langer gewenste verbeteringen maken.”

Uitnodigend en innovatief

Juryvoorzitter prof. Karel Luyben benadrukte namens de andere juryleden de praktische kanten en de mooie interface van DINED. “De dataset is

heel praktisch opgebouwd waardoor het maatschappelijk nut groot is. Het innovatieve karakter van de set zal veel wetenschappers uitnodigen voor verder onderzoek. Door veel aandacht te schenken aan zowel de verzameling, beschrijving en het beschikbaar stellen van de set, is deze voor een groot publiek zeer toegankelijk.”

Meer datasets

Naast de winnaars waren vier andere datasets genomineerd. Al zijn deze datasets niet bekroond, alle inzendingen waren mooie en inspirerende voorbeelden van het toegankelijk maken en delen van onderzoeksdata. Initiatiefnemer Research Data Netherlands is voornemens over twee jaar de prijs weer uit te reiken.

liederenbank.nl

dined.io.tudelft.nl

researchdata.nl

KB ResearchLab digitale speeltuin voor wetenschappers

Online aan de slag met data en tools

Wetenschappers kunnen in het KB ResearchLab zelf experimenteren met data van de KB en open-source tools. *Inge Angevaere*

De Koninklijke Bibliotheek (KB) heeft ruim 9 miljoen krantenpagina's, 1,5 miljoen tijdschriftpagina's en zo'n 90.000 andere publicaties gedigitaliseerd. Al die bestanden staan online op delpher.nl, waar iedere Nederlander ze kan doorzoeken. Maar onderzoekers willen vaak

meer dan alleen maar zoeken. Daarom biedt de KB sinds kort een omgeving aan waar wetenschappers met data en tools kunnen experimenteren: het KB ResearchLab.

Resultaten clusteren

Willem Jan Faber, manager van het Lab bij de KB: “Niet alleen de wetenschappers maar ook wijzelf staan nog aan het begin van een ontdekkingsstocht in Big Data. Om het maximale rendement uit die data te halen, werken we samen met onder-

zoekers. We bieden tools en data aan en nodigen wetenschappers uit om ermee te gaan werken.” Ontwikkelt de KB die tools ook? Faber: “De KB is natuurlijk vooral een leverancier van content. Maar uit onze samenwerking met onderzoekers ontstaan soms ook tools die de content beter doorzoekbaar maken of die helpen om resultaten te clusteren. Die kunnen we dan weer aanbieden aan andere wetenschappers. De samenwerking met onderzoekers is trouwens niet helemaal nieuw. De

afgelopen jaren hebben we al wat ervaring opgedaan, bijvoorbeeld met taalkundige Marc van Oostendorp (zie E-data februari 2014, red), met wie we een ‘scandeermachine’ bouwden. Samen met het Instituut voor Nederlandse Lexicologie hebben we gewerkt aan een tool voor het taggen van ‘named entities’ (eigenamen). Nu zijn we zo ver dat we de mogelijkheden van ons lab breder kunnen aanbieden, aan alle onderzoekers die willen experimenteren.”

VERVOLG OP PAGINA 2



**E-DATA &
RESEARCH**

Jaargang 9 | nummer 1

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

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

INHOUD

3 **Twitter: hype of here to stay voor de wetenschap?**

4 **Deze datasets zijn ook sinds kort beschikbaar**

5 **CLARIAH werkt aan een digitale infrastructuur**



6 **Lokale kennis als data mee in onderzoek**

6 **Deuren open voor virtueel onderzoek**

7 **RDA-week in beeld**

7 **KB: toegang tot ieder wetenschappelijk artikel**

8 **Hay Kranen: 'Wikipedia goudmijn voor onderzoek'**

8 **De bomen en het bos binnen de e-humanities**



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