De Nederlandse Dataprijs gaat naar 17e-eeuwse krantenset

20-11-2020 09:03



De Nederlandse Dataprijs 2020 in de categorie sociale en geesteswetenschappen is toegekend aan de 17e-eeuwse krantenset die door honderden vrijwilligers is getranscribeerd binnen een samenwerkingsproject tussen het Instituut voor de Nederlandse Taal (INT), de Koninklijke Bibliotheek, de nationale bibliotheek van Nederland (KB) en het Meertens Instituut onder leiding van Nicoline van der Sijs.

De dataset bestaat uit de oudste Nederlandstalige kranten, gepubliceerd tussen 1618 en 1700. Door de ouderdom van de tekst en de veelheid aan lettertypes konden deze kranten niet door de computer worden gelezen, maar ruim 200 vrijwilligers brachten uitkomst! Zij tikten de circa 120.000 artikelen over binnen een crowdsourcingsproject dat was opgezet door INT, KB en Meertens Instituut en dat onder leiding stond van Nicoline van der Sijs. Het resultaat is een betrouwbaar full-text-doorzoekbaar onderzoekscorpus van circa 20 miljoen woorden, dat sinds mei op Delpher staat: een online zoeksysteem dat toegang geeft tot ruim 120 miljoen pagina’s uit Nederlandse boeken, kranten en tijdschriften uit de 17e tot de 20e eeuw.

De Nederlandse Dataprijs bestaat uit een sculptuur en een prijzenbedrag van €5.000,-. INT en KB willen dit bedrag gebruiken om de data via een vervolgproject nóg verder te ontsluiten. Aan de vrijwilligers zal worden gevraagd alle Amsterdamse locaties (zoals straten, kroegen, markten) die genoemd worden in de zeventiende-eeuwse krantenartikelen, te identificeren en te koppelen aan een plattegrond van Amsterdam. Het resultaat zal een digitale kaart zijn waarop te lezen is wat waar gebeurde in zeventiende-eeuws Amsterdam en wat de toenmalige hotspots waren.

Projectleider Nicoline van der Sijs ziet de Nederlandse Dataprijs als ‘een geweldige erkenning van het maatschappelijk belang en de waarde van citizen science als bijdrage aan wetenschappelijk onderzoek’.

David Huijzer