

# 21e Colloquium Neerlandicum      Nijmegen, Nederland • 22-26 augustus 2022

---

## Submission 79

### DBNL als goudmijn voor de digitale neerlandistiek

Ryanne Keltjens<sup>1</sup>, Andreas Cranenburgh<sup>4</sup>, Marieke Erp<sup>2</sup>, Inger Leemans<sup>3</sup>

<sup>1</sup> KB, nationale bibliotheek, Den Haag, Nederland

<sup>2</sup> KNAW humanities cluster, Amsterdam, Nederland

<sup>3</sup> Vrije Universiteit Amsterdam, Amsterdam, Nederland

<sup>4</sup> Rijksuniversiteit Groningen, Groningen, Nederland

*Digital Humanities* klinkt voor velen als een vak apart. Maar ook binnen de neerlandistiek worden deze innovatieve methoden steeds belangrijker. Bovendien zijn ze lang niet zo hoogdrempelig als het lijkt; zelfs studenten kunnen digitale bibliotheken makkelijk doorzoeken en zo nieuwe onderzoeksvragen verkennen en beantwoorden. Nieuwe methoden zoals *distant reading* leiden niet alleen tot nieuwe inzichten over de Nederlandse taal en cultuur, maar brengen vaak ook een uitbreiding van onderzoeksobject, methoden en perspectieven met zich mee, waardoor disciplinaire grenzen verlegd worden. Voor de digitale neerlandistiek is de collectie van de Digitale Bibliotheek voor de Nederlandse Letteren (DBNL) van groot belang. Werken uit de DBNL-collectie worden niet alleen bestudeerd voor kwalitatief onderzoek, maar ook als dataset gebruikt voor kwantitatieve analyses binnen de *digital humanities*. In deze panelsessie verkennen we aan de hand van drie projecten die gebruik maken van DBNL-data hoe kwantitatieve analyses van digitale corpora de internationale neerlandistiek verder kunnen brengen.

Zijn romans uit de literaire canon te onderscheiden van andere romans op basis van tekstkenmerken? Andreas van Cranenburgh heeft een dataset met tekstkenmerken en metadata van DBNL-romans verzameld en beschikbaar gemaakt. DBNL houdt verwijzingen in secundaire literatuur naar primaire teksten systematisch bij. Deze verwijzingen heeft Andreas gebruikt als *canoniciteits-indicator*: hoe vaker een roman besproken wordt door critici en letterkundigen, hoe aannemelijker het is dat deze tot de canon behoort. Voor zijn eerste resultaten was de aanname dat elke roman met minstens één verwijzing tot de canon behoort. Een logistische regressie op basis van woordfrequenties weet 72% van de romans correct te identificeren als canonic of niet. Deze resultaten zijn in lijn met de hypothese dat er een verband is tussen tekstkenmerken en canoniciteit.

Wat is de geur van de Nederlandse letterkunde? Welke rol spelen geuren in de Nederlandse literatuur? Het Europese Horizon-2020-project Odeuropa ontwikkelt *sensory mining* technieken om geurerfgoed op te sporen en terug ten neuze te brengen. In Nederland is nog maar weinig gedaan met geurgeschiedenis. Dat is jammer want we beschikken over rijke historische teksten en beeldcorpora, zoals DBNL. In het Odeuropa-project (dat zich uitstrekt over zeven taalgebieden) maken we dan ook dankbaar gebruik van de hoog-kwalitatieve data van DBNL. We speuren naar olfactorische informatie: naar vermeldingen van geurstoffen, waarderingen, sentimenten, gebruiken en 'neuzen'. In deze presentatie leggen Marieke van Erp en Inger Leemans uit wat de doelstellingen zijn van het Odeuropa-project, en met welke methoden zij het rijke geurverleden van Nederland proberen in kaart te brengen.

[samenvatting spreker 3 volgt nog]