



K O N I N K L I J K E N E D E R L A N D S E
A K A D E M I E V A N W E T E N S C H A P P E N

TS Tools: Nederlab voor tijdschriftonderzoek

van der Sijs, N.

published in

TS. Tijdschrift voor Tijdschriftstudies
2015

document version

Publisher's PDF, also known as Version of record

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

van der Sijs, N. (2015). TS Tools: Nederlab voor tijdschriftonderzoek. *TS. Tijdschrift voor Tijdschriftstudies*, 37, 53-62.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl

URN:NBN:NL:UI:10-1-114184. TS •> #37, June 2015, p. 53-62

Content is licensed under a Creative Commons Attribution 3.0 License. - © Nicoline van der Sijs

Publisher: www.uopenjournals.org. Website: www.tijdschriftstudies.nl

TS Tools: Nederlab voor tijdschriftonderzoek

NICOLINE VAN DER SIJS

post@nicolinevdsijs.nl

ABSTRACT

March 2015 saw the launch of the infrastructure Nederlab, a laboratory for research on the patterns of change in the Dutch language and culture. In this section Nicoline van der Sijs, coordinator of Nederlab, shows how Nederlab can be used for research into periodical media.

KEY WORDS

Digitization, periodical studies, Nederlab, e-humanities

TER INLEIDING

Het leven van onderzoekers in het algemeen en van tijdschriftonderzoekers in het bijzonder is de laatste tien jaar een stuk minder avontuurlijk geworden. Vroeger moest je barre tochten door Nederland ondernemen naar ijskoude, in donkere kelders gehuisveste archieven waar je onder het toeziend oog van een strenge bibliothecaris door langzaam uit elkaar vallende pagina's van oude tijdschriften en kranten moest bladeren op zoek naar een lang vergeten of onopgemerkt gebleven pareltje waarmee je je cv kon uitbreiden. Tegenwoordig hoef je thuis of op je werk maar je computer aan te zetten om een overvloed aan oude en jonge tijdschriftnummers systematisch digitaal te kunnen doorvlooiën. En niet alleen jij kan dat, ook je wetenschappelijke collega's, zodat je niet langer weg kunt komen met een artikel dat niet meer biedt dan een anekdotisch feitje of een grappig citaat van een bekende of lang vergeten auteur.

Er is dus digitaal materiaal in overvloed, maar dat levert, behalve nieuwe mogelijkheden, vooral ook veel nieuwe problemen op. Als je het lezen uitbesteedt aan de computer, hoe kun je er dan zeker van zijn dat de computer vindt wat er is – niet meer, maar ook niet minder? Hoe kun je dat zoeken verbeteren? Hoe kun je je weg vinden in het grote aantal resultaten dat de computer je voorschotelt? En hoe kun je daaraan duiding, betekenis geven?

Om onderzoekers met deze en nog veel meer vragen te helpen is op 1 januari 2013 het NWO-groot project Nederlab van start gegaan. Het doel van Nederlab is wetenschappers een onderzoeksinstrumentarium te bieden, en een laboratorium waarin zij vrijelijk kunnen experimenteren met data en tools. Op 13 maart van dit jaar is de betaversie van [Nederlab](#) gelanceerd. Hieronder zal ik wat algemene informatie over Nederlab geven, maar vooral laten zien wat tijdschriftonderzoekers aan Nederlab kunnen hebben. Ter illustratie vermeld ik enkele kleine casussen.

ONDER ÉÉN DAK

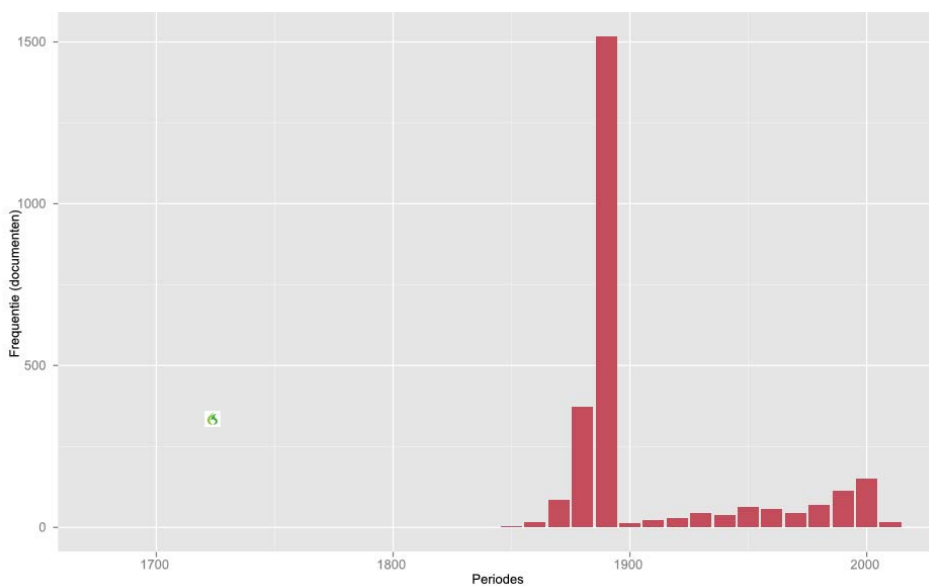
Momenteel worden veel gedigitaliseerde collecties van boeken, tijdschriften en kranten op verschillende plaatsen aangeboden: tijdschriften en seriewerken kun je bijvoorbeeld vinden bij de DBNL, de Koninklijke Bibliotheek, EDBO en Google Books. Delpher biedt een toegangspoort op deze collecties, en maakt het mogelijk ze tegelijkertijd te doorzoeken. Maar de zoekresultaten worden vervolgens per collectie aangeboden en ze worden niet bij elkaar opgeteld. Bovendien zijn ook buiten Delpher kranten, tijdschriften of seriewerken te vinden, bijvoorbeeld de [Startpagina historische kranten](#), [Staten-Generaal Digitaal](#), [Elsevier Erfgoed](#), en tijdschriftjaargangen van o.a. Knack en Weekend Knack binnen [OpenSonar](#).

Het doel van Nederlab gaat verder dan Delpher: Nederlab wil zoveel mogelijk collecties, niet alleen die van de Koninklijke Bibliotheek maar ook bijvoorbeeld die van Huygens ING, Meertens Instituut, Instituut voor Nederlandse Lexicologie en universiteitsbibliotheken, gezamenlijk doorzoekbaar maken via één portaal. En de zoekresultaten kunnen integraal worden geanalyseerd en gevisualiseerd. Op het verlanglijstje van Nederlab staan ongeveer 50 collecties. Zover is het nu nog niet: momenteel bevat Nederlab drie omvangrijke collecties: DBNL, EDBO en de historische kranten van de Koninklijke Bibliotheek. Die drie collecties zijn wel gezamenlijk doorzoekbaar gemaakt.

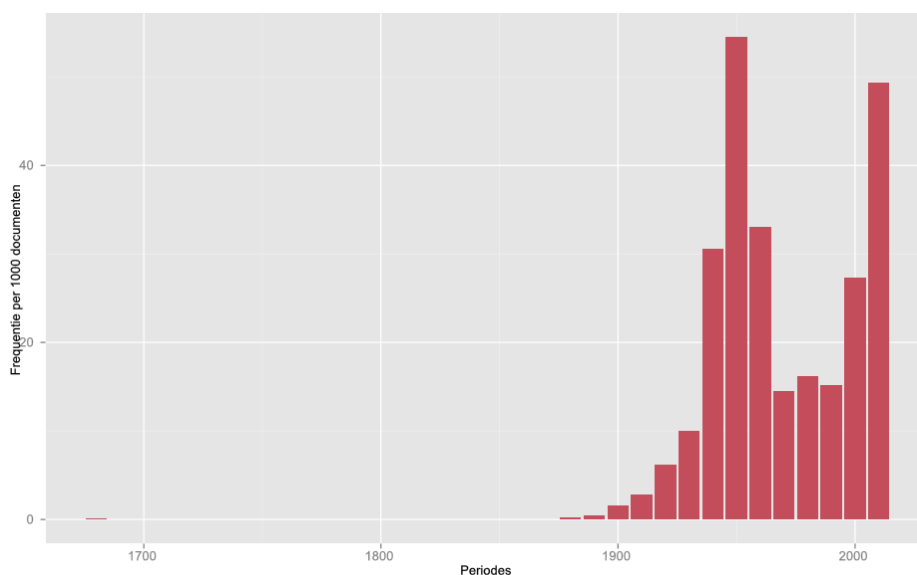
Voor statistisch onderzoek is het belangrijk om goede referentiewaarden te hebben. Daarom biedt Nederlab op de homepage www.nederlab.nl/onderzoeksportaal direct inzicht in de omvang van de verschillende collecties, uitgezet op een tijdsas. Op de homepage wordt ook aangegeven wanneer er nieuwe collecties aan Nederlab zijn toegevoegd. Onderzoekers die zijn ingelogd in Nederlab krijgen bovendien een overzicht te zien van het totale aantal documenten van alle collecties samen, en hoe die zijn verdeeld over de tijd. Een dergelijk overzicht is essentieel om een resultatenlijst te duiden. Iedere onderzoeker kan inloggen met het e-mailadres van zijn instelling. Belangstellenden die niet verbonden zijn aan een wetenschappelijke instelling kunnen alleen gebruikmaken van de algemene functies van Nederlab. Willen ze meer, dan kunnen ze Nederlab verzoeken om een speciale inlogcode.

Onderzoekers kunnen, of ze nu zijn ingelogd of niet, een visueel overzicht opvragen van zoekresultaten, als hulpmiddel voor de interpretatie van de resultaten. Als je bijvoorbeeld in de tekst het woord *detective* zoekt, krijg je een plaatje te zien van de verbreiding van het woord *detective* door de tijd heen (afbeelding 1). Op het moment dat ik deze zoekopdracht doe, 27 mei 2015, geeft de periode 1890-1899 een enorme piek te

zien. Hoe is dat te verklaren? Was er op dat moment een gebeurtenis waardoor detectives in het nieuws kwamen, was er een prijs uitgereikt voor een detectiveverhaal of, waarschijnlijker in deze periode, was er een pamflet verschenen dat de jeugd waarschuwde tegen de verderfelijke invloed van de Engelse detectives? Niets van dit al, zo blijkt als je kijkt naar de opbouw van het complete corpus: de enige reden dat er een piek is in de periode 1890-1899, is dat er in die periode veel kranten beschikbaar zijn, waardoor de opbouw van het complete corpus vertekend is. Als je ingelogd bent, kun je binnen Nederlab het corpus van je zoekresultaten bewaren. En dan krijg je veel meer visualiseringsmogelijkheden (die zijn ontwikkeld door Erwin Komen): je kunt nu bijvoorbeeld ook de relatieve resultaten zichtbaar maken en niet alleen de absolute. Uit de relatieve resultaten (afbeelding 2) blijkt dan dat de piek in een heel andere periode ligt, namelijk in 1950-1959, de periode van schrijvers als Havank, Maigret en Fleming.



Afb. 1: De absolute documentfrequentie van het woord *detective* in Nederlab



Afb. 2: De relatieve documentfrequentie van het woord *detective* in Nederlab

Momenteel krijgen gebruikers in Nederlab alleen documentfrequentie te zien, dus in welke documenten een bepaalde term voorkomt, maar niet hoe vaak binnen dat document. In een volgende versie van Nederlab willen we de mogelijkheid bieden om ook termfrequenties op te vragen.

Overigens blijft de alertheid van de onderzoeker essentieel – dit ter geruststelling van degenen die bang zijn dat de nieuwe digitale mogelijkheden onderzoekers overbodig maken. Want wie denkt dat hij op basis van afbeelding 2 de geschiedenis van de detective in het Nederlands kan schrijven (eerste detective rond 1890, piek in 1960), komt bedrogen uit. Als je namelijk de snippets van de zoekresultaten leest, blijkt dat in de periode tot 1890 uitsluitend sprake is van menselijke detectives – geheime politieagenten dus. In de jaren 1890-1899 komen *detectiveromans*, *detectivestories* en *detectiveverhalen* op (aldus genoemd), en pas in de loop van de 20ste eeuw worden deze samenstellingen verkort tot de benaming *detective*.

VERRIJKINGEN VAN DE DATA

Weinig verrassend is dat je in Nederlab in teksten kunt zoeken, met wildcards en wel of niet hoofdlettergevoelig. Handig is de mogelijkheid om woordvarianten mee te nemen in je zoekopdracht, zonder dat je hoeft te weten welke spellingsvarianten er allemaal bestaan – een notoir probleem in het analyseren van oude teksten. Die zoekmogelijkheid met woordvarianten is gebaseerd op het historisch computationeel lexicon van het Instituut voor Nederlandse Lexicologie, en ook Delpher biedt deze mogelijkheid. Dit historische lexicon is nog in bewerking. Enerzijds ontbreken er nog gegevens: als je bijvoorbeeld zoekt op de woordvarianten van *cadeau*, wordt alleen *cadeaux* gepresenteerd, maar *cadeaus*, *kado*, *kado's*, *kadoos* ontbreken, en er wordt geen enkele verkleinvorm (*cadeautje*, *kadootje* etc.) meegenomen. In andere gevallen worden er juist te veel varianten getoond: als je de woordvarianten van *paard* zoekt, krijg je de verwachte vormen *peerd* en *paerdt*, maar ook *paren*, *paar* en *gepaard* – die niets met het dier te maken hebben. Binnen Nederlab willen wij in een volgende versie de mogelijkheid bieden dergelijke vormvarianten weg te klikken uit de zoekresultaten, zodat onderzoekers de resultaten kunnen inperken tot precies die waarnaar ze op zoek zijn.

Om het mogelijk te maken alle teksten gelijktijdig te doorzoeken, zetten we ze allemaal om naar een uniform formaat.¹ Via de parser [Frog](#) voegen we lemmatiseringen toe en informatie over de woordsoort, zodat het bijvoorbeeld mogelijk wordt alleen *arm* in de betekenis ‘behoefstig’ te zoeken, en het lichaamsdeel *arm* op voorhand uit de zoekresultaten weg te filteren, opnieuw met als doel de onderzoeker te helpen bij het inperken van de zoekresultaten. De mogelijkheid te zoeken op woordsoort is overigens nog niet geïmplementeerd. De teksten die met optische tekenherkenning zijn gelezen, voorzien we tot slot via het programma TICCL (Text-Induced Corpus Clean-up), ontworpen door Martin Reynaert, van een extra, gecorrigeerde laag, zodat veel van de leesfouten automatisch worden gecorrigeerd.

¹ Voor de technici: het betreft het formaat [FoLiA](#).

DE KRACHT VAN METADATA

Voor tijdschriftonderzoekers ligt de kracht van Nederlab in de metadata: de auteurs- en titelgegevens. Nederlab heeft een eigen auteursthesaurus en titelthesaurus ontwikkeld. Wanneer nieuwe collecties worden binnengehaald, worden de auteurs uit die collecties semi-automatisch gelinkt met de auteursnamen die al binnen Nederlab beschikbaar zijn. De redactie van Nederlab heeft hiervoor in samenwerking met de technische afdeling van het Meertens Instituut een zogenaamde harmonisatietool ontwikkeld, die voornaam, achternaam, geboortjaar en sterfjaar van een auteur uit een nieuwe collectie automatisch vergelijkt met de Nederlab-auteursgegevens: is er een match, dan worden ze aan elkaar gekoppeld; is er twijfel, dan bekijkt een klein redactieteam onder leiding van René van Stipriaan handmatig of een koppeling tot stand gebracht kan worden.

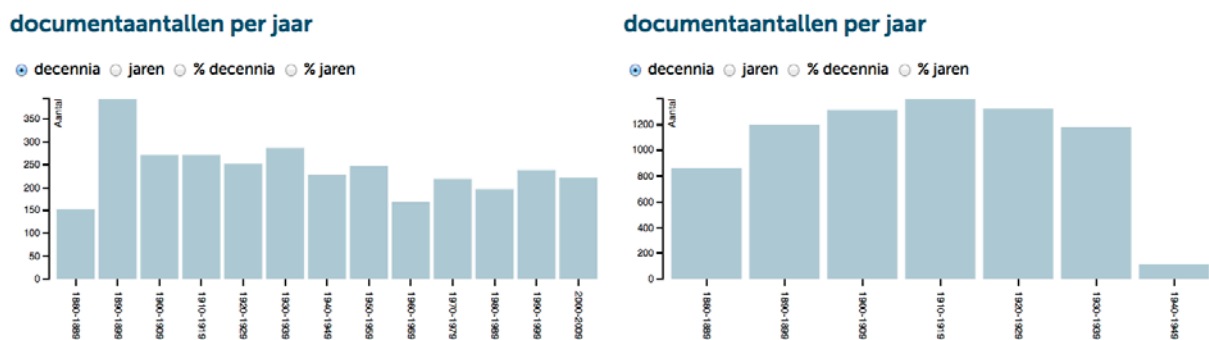
Het resultaat hiervan biedt ongekende nieuwe mogelijkheden voor het gericht zoeken en analyseren van de gegevens. Je kunt bijvoorbeeld alle werken van één of meer auteurs tegelijkertijd voor onderzoek selecteren, ook als die werken afkomstig zijn uit verschillende collecties. Die mogelijkheid wordt nergens elders geboden. In de toekomst kun je ook werken die identiek zijn (ongewijzigde herdrukken) uitsluiten door te kiezen voor: Doublures uitsluiten. En je kunt ook alle artikelen uit één tijdschrift opvragen door in Titelgegevens te zoeken op Koepeltitel. Als je bijvoorbeeld in Koepeltitel het oudste Nederlandstalige tijdschrift, *De boekzaal van Europe*, opzoekt, zie je in één oogopslag dat dit tijdschrift heeft gelopen van 1692 tot 1702 en in totaal 114 artikelen bevatte (afbeelding 3).

The screenshot shows the Nederlab search interface. At the top, the logo 'nederlab' is displayed with navigation links: 'mijn nederlab', 'zoeken', 'over nederlab', 'help', and 'uitloggen'. Below the logo, there are three search sections: 'zoeken in tekst', 'zoeken in titelgegevens', and 'zoeken in auteursgegevens'. The 'zoeken in titelgegevens' section is active, showing a search for 'De boekzaal van Europe (1692-1702)'. The search results show 114 results. Below the search bar, there are filters for 'samenstelling' and 'beperk tot genre(s)'. The 'samenstelling' filter is set to 'zelfstandige titels', 'onzezelfstandige titels', and 'koepeltitels'. The 'beperk tot genre(s)' filter is set to 'drama' and 'jeugdletteratuur'. A red button 'bewaren als corpus' is visible, along with links for 'lijst van bronnen' and 'visueel overzicht'.

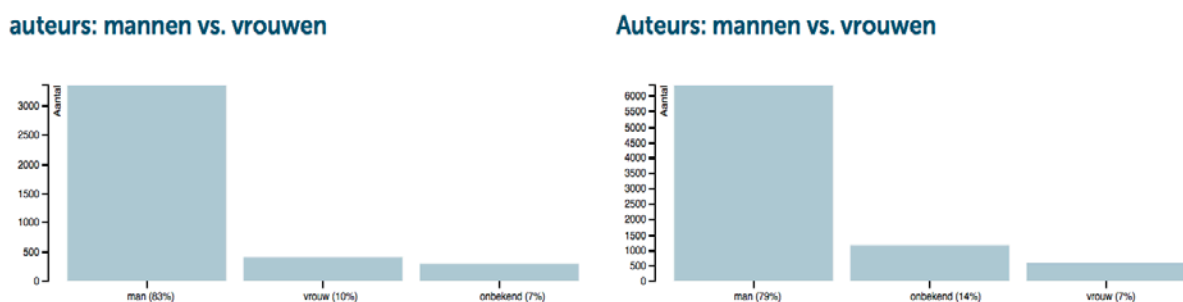
Afb. 3: Zoeken op Koepeltitel in titelgegevens

Maar de mogelijkheden voor tijdschriftonderzoekers reiken nog veel verder. Met steun van CLARIN-NL hebben René van Stipriaan en Erik Tjong Kim Sang voor Nederlab een vergelijkingsmodule gebouwd die het mogelijk maakt de metadagegevens van twee verschillende corpora, bijvoorbeeld jaargangen van twee tijdschriften, met elkaar te vergelijken. Als voorbeeld heb ik alle jaargangen genomen van het *Tijdschrift voor Nederlandse Taal- en Letterkunde* (TNTL), dat begon in 1881, en de jaargangen van *De Gids* vanaf 1881 (*De Gids* begon in 1837). De onderliggende data zijn momenteel nog in bewerking (zo ontbreken veel jaargangen van *De Gids* na 1940 vanwege copyrightrestricties), dus we kunnen nog geen harde conclusies verbinden aan de vergelijkingen, maar de afbeeldingen 4 t/m 10 tonen wel wat er allemaal mogelijk is.

We kunnen bijvoorbeeld een vergelijking maken tussen het aantal documenten (lees: artikelen) per jaar voor ieder tijdschrift (afbeelding 4). *De Gids* blijkt dan een stuk omvangrijker te zijn, in alle periodes, dan *TNTL*, maar dit laatste tijdschrift heeft een interessante piek in de jaren 1890-1899, en na een behoorlijke dip in de jaren 1960-1969 blijft het aantal artikelen in de daarop volgende jaren vrij stabiel. Een onderzoeker heeft nu alle aanleiding om eens beter naar de betreffende jaargangen te kijken: wat is hier aan de hand, wijziging in het beleid of stom toeval? Vervolgens kan ook bekeken worden of er sprake is van een breder patroon, en of er specifieke externe factoren werken, zoals papierschaarste, toenemende of afnemende concurrentie enz.



Afb. 4: Een vergelijking tussen *TNTL* (links) en *De Gids* (rechts): het aantal artikelen

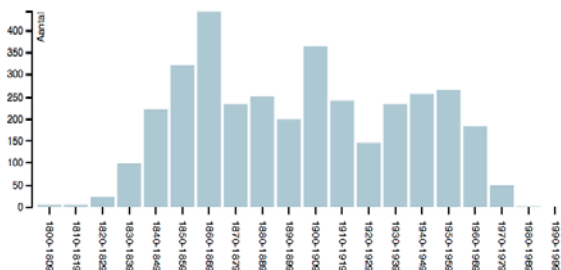


Afb. 5: Een vergelijking tussen *TNTL* (links) en *De Gids* (rechts): het geslacht van de auteurs

Verder kunnen we het aantal vrouwen bekijken dat een bijdrage aan de tijdschriften heeft geleverd (afbeelding 5). Hier zien we weinig verschil tussen de tijdschriften. (Onbekend betekent dat er geen sekse-informatie over de betreffende auteurs aanwezig is in de auteursdatabase.) Ook kunnen we de geboortejaren en overlijdensjaren van auteurs met elkaar vergelijken (afbeeldingen 6 en 7). Ik onthoud me hier van exegese, maar het is duidelijk dat in sommige jaren meer auteurs zijn geboren dan in andere, wat betekent dat in sommige periodes vooral tijdgenoten in het tijdschrift schreven; is hier wellicht sprake van vriendjespolitiek? Of is het tijdschrift volgeschreven door telkens hetzelfde kleine groepje auteurs?

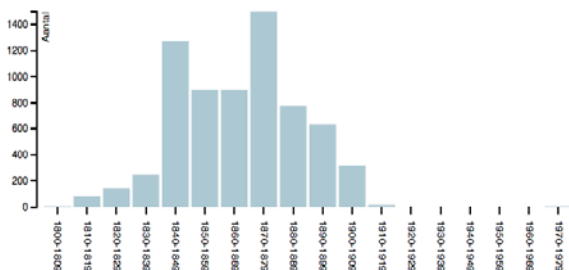
geboortejaren auteurs

decennia jaren



geboortejaren auteurs

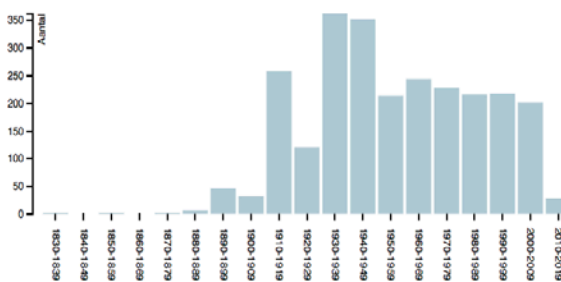
decennia jaren



Afb. 6: Een vergelijking tussen *TNTL* (links) en *De Gids* (rechts): de geboortejaren

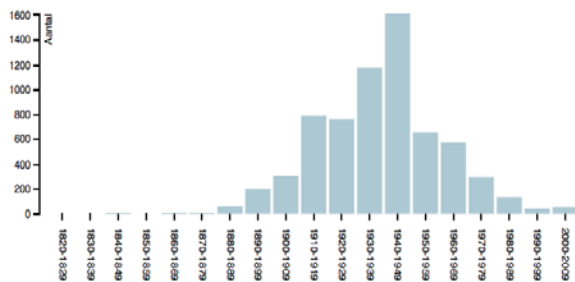
overlijdensjaren auteurs

decennia jaren



overlijdensjaren auteurs

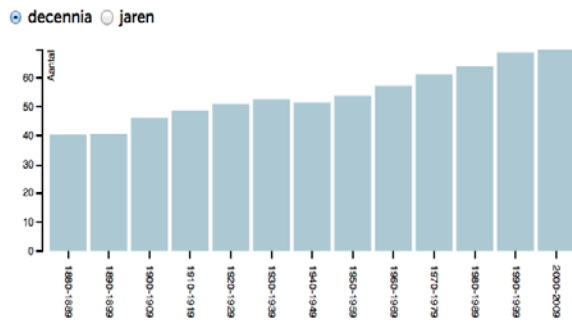
decennia jaren



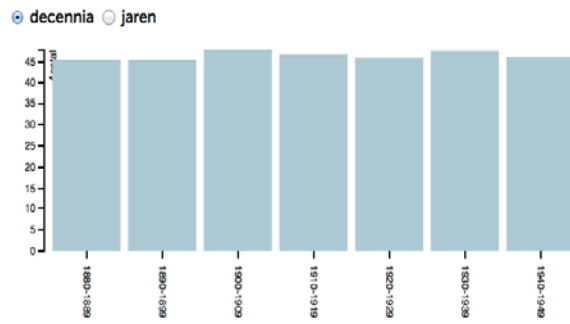
Afb. 7: Een vergelijking tussen *TNTL* (links) en *De Gids* (rechts): de overlijdensjaren

Minstens zo interessant is het overzicht van de gemiddelde leeftijd (afbeelding 8), waaruit blijkt dat de auteurs van *TNTL* almaar ouder worden en hun gemiddelde leeftijd momenteel rond de 65 jaar ligt. Bij *De Gids* is voor de periode 1880-1940 de gemiddelde leeftijd redelijk constant en rond de 45, daar is kennelijk sprake van continue verjonging. Tot slot zeggen de geboorteprovincies en overlijdensprovincies (afbeeldingen 9 en 10) iets over de herkomst van auteurs. Het blijkt dat de auteurs van beide tijdschriften vooral geboren zijn in Noord-Holland, terwijl als sterfprovincie Zuid-Holland het hoogst scoort. Is hier sprake van een trek van Noord-Holland naar Zuid-Holland? Er zijn in de afbeeldingen meer verschillen te zien, maar daarop ga ik hier niet in detail in.

gemiddelde leeftijden auteurs

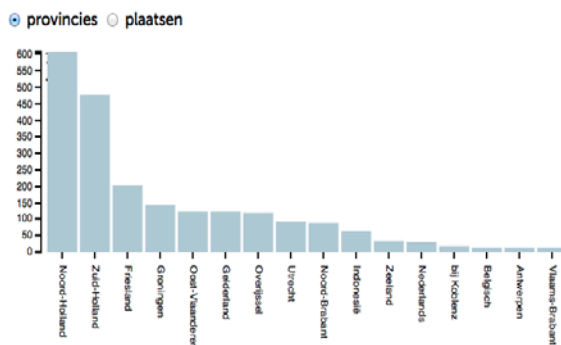


gemiddelde leeftijden auteurs

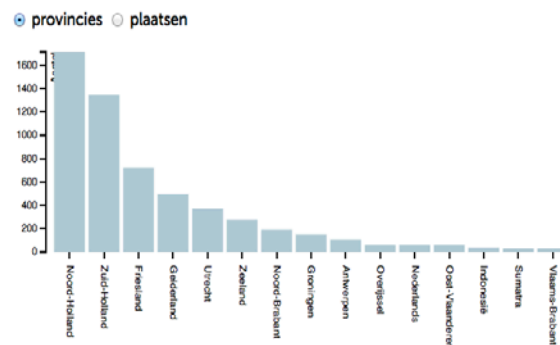


Afb. 8: Een vergelijking tussen *TNLT* (links) en *De Gids* (rechts): de gemiddelde leeftijd

geboorteprovincies auteurs

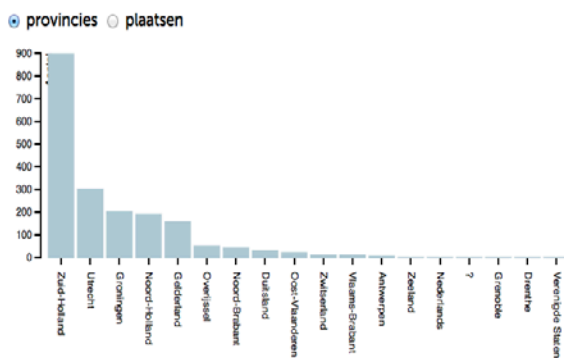


geboorteprovincies auteurs

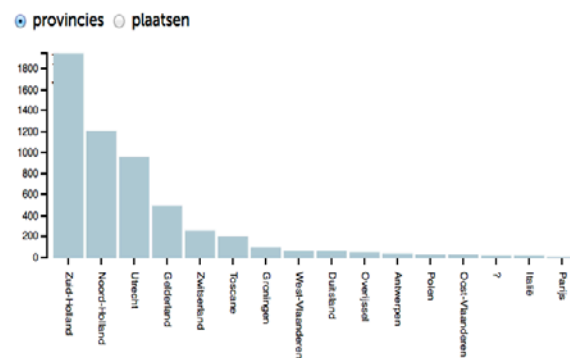


Afb. 9: Een vergelijking tussen *TNLT* (links) en *De Gids* (rechts): de geboorteprovincie

overlijdensprovincies auteurs



overlijdensprovincies auteurs



Afb. 10: Een vergelijking tussen *TNLT* (links) en *De Gids* (rechts): de overlijdensprovincie

DE TACHTIGERS

Nederlab maakt het dus mogelijk een beeld, en een tijdsbeeld, te krijgen van de auteurs die gezamenlijk de inhoud van een tijdschrift hebben bepaald. Daarnaast kan via Nederlab ook de inhoud van de tijdschriften zelf worden onderzocht. En dat maakt het bijvoorbeeld mogelijk een vraag te beantwoorden als: in de annalen van welk tijdschrift is de naam voor de Tachtigers ontstaan? Was het een geuzennaam die opkwam in *De Nieuwe Gids*, het tijdschrift dat de leden van de literaire stroming in 1885 oprichtten om zich af

te zetten tegen de in hun ogen moralistische en retorische literatuur die hun voorgangers sinds 1843 publiceerden in *De Gids*? Of was het juist de oude garde die in *De Gids* het aanstormende talent denigrerend vernoemde naar de periode waarin zij publiceerden?

Als we binnen Nederlab zoeken naar ‘tachtigers’ vanaf 1880 (een eerdere periode kan om voor de hand liggende redenen worden uitgesloten), vinden we als eerste hit *Grassprietjes* van Frederik van Eeden uit 1885. Bingo, kennelijk is het een geuzennaam geweest! Maar die conclusie is veel te snel: als je het complete citaat nazoekt, blijkt dat er sprake is van een editie van *Grassprietjes* uit 1984 en dat de benaming Tachtigers staat in het Nawoord van de toenmalige redacteur. Om dit soort vervuilde resultaten te voorkomen, willen we in de toekomst binnen Nederlab zoveel mogelijk automatisch oorspronkelijke teksten en redactieteksten (zoals later toegevoegde noten, voorwoorden en nawoorden) van elkaar scheiden.

Als we verder zoeken naar de eerste echte vermelding van de benaming Tachtigers blijkt die te staan in *De Gids* uit 1903, het tijdschrift dus waartegen de Tachtigers zich zo hevig verzetten. Het schrijversechtpaar Carel en Margot Scharten-Antink begint een recensie van het toneelstuk *Jacoba van Beijeren* van Albert Verwey met de karakterisering: ‘Albert Verwey, de rustigste, de gespierdste van dat geniaal, maar zoekend en wankel geslacht der tachtigers’.² Waren de tegenstanders dus verantwoordelijk voor de naam van de Tachtigers? Toch niet, blijkt als je nog wat dieper in de data duikt. Want in het jaar dat het tijdperk eindigde, in 1890, vat de redactie van *De Nieuwe Gids* in de rubriek ‘Literaire kroniek’ hun credo samen: ‘Wij lieden van ’80 [...] wij geven niet meer om fraaiheid van frase, om eruditie of vormlijkheid, wij zijn zoo natuurlijk en trouwhartig en kinderlijk in ons voelen en verbeelden, als goede artiesten dat altijd zijn geweest. Wij trachten in alles te wezen direct.’³ In datzelfde tijdschrift is vier jaar later, in 1894, sprake van ‘de invloed der beweging van ’80’.⁴ Vanaf dat moment is de term ‘de (letterkundige, litteraire, dichterlijke) beweging van ’80’ algemeen in gebruik. Totdat deze in 1903 wordt verkort tot Tachtigers – de benaming waarmee de beweging de geschiedenis ingaat.

TER AFSLUITING

Ik hoop dat de voorbeelden hebben getoond wat Nederlab dankzij zijn fijnmazige meta-datastelsel kan betekenen voor tijdschriftonderzoek. Die metadata maken het mogelijk grote hoeveelheden zoekresultaten te ordenen, te analyseren en te visualiseren. Ook bieden de Nederlab-metadata de mogelijkheid om bijvoorbeeld de omvang van verschillende tijdschriften met elkaar vergelijken, de netwerken van auteurs rond literaire en taalkundige tijdschriften in beeld te krijgen, te beoordelen of een tijdschrift verjonging representeert, en te bekijken in hoeverre vrouwen oprukken in tijdschriften die van oudsher mannenbolwerken waren.

² C. en M. Scharten-Antink, ‘Driemaandelijksch overzicht der Nederlandsche letteren II. De dichtkunst in 1903.’ *De Gids* 67:4, 1903, p. 108-161 (153).

³ ‘Literaire kroniek.’ *De Nieuwe Gids* 5, 1890, 420-425 (421).

⁴ P. Tideman, ‘Moderne kinderkamer.’ *De Nieuwe Gids* 9, 1894, 438-444 (441).

Daarmee zijn nog lang niet alle mogelijkheden uitgeput: als er meer specifieke metadata beschikbaar komen, neemt het aantal vergelijkingspunten toe. Zo is het leerzaam om te bekijken hoe groot de bijdragen van katholieken en joden aan tijdschriften was in de 19de eeuw, welke opleidingen auteurs door de tijd heen hebben gevolgd, en welke beroepen auteurs en recensenten uitoefenden om wat bij te verdienen. Het betreft hier overigens gegevens die Nederlab op dit ogenblik niet biedt, en ook niet op korte termijn zal gaan bieden. Uiteindelijk zullen onderzoekers van verschillende specialisatie binnen specifieke projecten dergelijke gegevens moeten aanbrengen. De reden daarvoor is eenvoudig: bepalen welk geloof iemand belijdt, is een delicate zaak, en vraagt de nodige expertise en ook wetenschappelijke voorbehouwen; hetzelfde geldt voor het bepalen van het opleidingsniveau en de voornaamste broodwinning – ook hierbij komt veel interpretatie kijken.

Er zijn dus meer metadata nodig. En er zijn ook véél meer data nodig! Hoe interessant zou het niet zijn om deftige literaire tijdschriften te kunnen vergelijken met modetijdschriften, de *Libelle*, een omroepblad of de *Allerhande*. Maar van dergelijk wegwerpdrukwerk is maar een heel klein deel bewaard gebleven en vooralsnog is dat niet gedigitaliseerd.

Ik begon met de opmerking dat het leven van onderzoekers de laatste decennia een stuk minder avontuurlijk is geworden, omdat ze niet meer de deur uit hoeven voor gegevens. Het onderzoek zélf is daarentegen nog nooit zo spannend en avontuurlijk geweest als nu: dankzij de digitalisering, en dankzij onderzoekslaboratoria als Nederlab, komen allerlei nieuwe onderzoeksvragen in het vizier die voorheen alleen maar met onmenselijk grote inspanningen te beantwoorden waren. We hopen dat tijdschriftonderzoekers eens een kijkje nemen bij Nederlab. En laat het ons gerust weten als u vragen of suggesties heeft: een berichtje naar post@nicolinevdsijs.nl volstaat.

•> NICOLINE VAN DER SIJS is coördinator van *Nederlab*, senior onderzoeker bij het *Meertens Instituut* en hoogleraar *Historische taalkunde van het Nederlands in de digitale wereld* aan de *Radboud Universiteit Nijmegen*.