

A LLE NEDERLANDSTALIGE LITERATUUR ONLINE

DE DIGITALE BIBLIOTHEEK
VOOR DE NEDERLANDSE LETTEREN

Vanaf de start in 2000 is de Digitale Bibliotheek voor de Nederlandse Letteren (DBNL) een succes. Iedere dag bekijken meer bezoekers een grotere selectie uit het almaar groeiende corpus literaire teksten in de databank. Helaas kan de DBNL niet beschikken over de gewenste capaciteit. Anders was de site al dichter bij zijn doel om de complete Nederlandstalige literatuur aan te bieden.

VAN ANTWERPS LIEDBOEK TOT SANNEKE VAN HASSEL

Hoe moet je in hemelsnaam aan werk van Filip De Pillecyn komen? De Vlaamse schrijver (1891-1962) heeft twee standbeelden en een eigen museum in zijn geboorteplaats Hamme. Maar in de boekhandel vind je, als je geluk hebt, alleen een winkel-dochter – een heruitgave van *Monsieur Hawarden* (1934), die alweer veertien jaar oud is. In antiquariaten liggen misschien vergeelde exemplaren van *De soldaat Johan* (1939) of *Mensen achter de dijk* (1946). Alleen op internet is meer te vinden. De Digitale Bibliotheek voor de Nederlandse Letteren biedt nagenoeg het complete werk van De Pillecyn aan. Op woordniveau doorzoekbaar.

En dat is lang niet alles. Verre van. Het *Antwerps liedboek* uit 1544. *Achilles en Polyxena* van P.C. Hooft. *Klaasje Zevenster* van Jacob van Lennep. *Pallieter* van Felix Timmermans. *Ijsregen* van Sanneke van Hassel. En nog duizenden en duizenden titels. Ze zijn allemaal van de eerste tot de laatste bladzijde te vinden op www.dbnl.org – niet zelden aangevuld met biografieën uit het *Letterkundig woordenboek voor Noord en Zuid* van K. ter Laan (1952) of *De Nederlandse en Vlaamse auteurs* van G.J. van Bork en P.J. Verkruijssse (1985), en secundaire artikelen uit literaire

MAARTEN DESSING

werd in 1974 geboren in Naaldwijk. Hij studeerde journalistiek aan de School voor Journalistiek. Was redacteur van *Boekblad*, schrijft sinds 2003 als freelancejournalist over boeken en het boekenvak voor onder meer *Knack*, *BOEK*, *HP/De Tijd*, *Boekblad*, *Bibliotheekblad* en *Taalschrift*. Was van 2009 tot 2012 jurylid van de AKO Literatuurprijs.
Adres: mdessing@yahoo.com.

tijdschriften. Ook de jaargangen van *Ons Erfdeel* sinds 1962 zijn er bijna integraal te lezen.

Als het aan de medewerkers van de DBNL had gelegen, bevatte de databank anno 2013 nog veel meer teksten, vertelt directeur Cees Klapwijk. Maar hij en zijn team van – inclusief hemzelf – zes medewerkers kunnen maar zo snel digitaliseren als er geld beschikbaar is. Dat blijft al jaren achter bij zijn wensen. “Al in 2005 zagen beleidsmakers in hoe succesvol de DBNL is”, zegt René van Stipriaan, tot 1 januari 2013 hoofdredacteur van de DBNL. “We hebben een groot publiek uit allerlei geledingen van de samenleving. We krijgen daar alle lof voor, ook vanuit de politiek, maar die goodwill vertaalt zich niet in een gezonde groei van onze bekostiging.”

De DBNL krijgt jaarlijks iets meer dan negenhonderdduizend euro van de Nederlandse Taalunie. Ondanks een steeds grotere database die up-to-date moet worden gehouden en een groeiend aantal diensten, is dat bedrag al meer dan vijf jaar niet meer verhoogd, aldus Van Stipriaan. Sterker: door de structurele bezuinigingen van de Nederlandse overheid is de subsidie sinds 2012 met vijf procent gekort. Daarnaast heeft de DBNL last van de crisis waardoor het – “al sinds 2008”, zegt Van Stipriaan – moeilijker is om projectgelden binnen te halen. Er waren jaren dat vijftig procent van het budget uit samenwerkingsopdrachten kwam, nu is dat naar schatting nog maar twintig tot dertig procent.

Het gevolg was dat de DBNL van een aantal medewerkers afscheid moest nemen. In 2010 telde de formatie nog tien mensen. “Door wat te herschikken kunnen we zo nu en dan tijdelijk weer iemand aannemen”, zegt Klapwijk, “maar eigenlijk zijn we

structureel met te weinig om de volgende stap te zetten. Een integralistische aanpak: alle middeleeuwse teksten, gehele oeuvres of clusters van tijdschriften – we zijn er al jaren klaar voor om over te gaan naar deze fase, maar voor een massale vorm van digitalisering én het zo goed mogelijk aanbieden van deze teksten aan de doelgroep zijn investeringen nodig. En die kunnen we juist niet doen.”

De aanstaande verhuizing naar de Koninklijke Bibliotheek in Den Haag – per 1 april 2013, is de planning – maakt vooralsnog geen verschil. De DBNL bespaart hiermee in de kosten voor huisvesting en andere faciliteiten, waardoor er theoretisch meer geld beschikbaar is voor de digitalisering. Maar daar staat tegenover dat Klapwijk nog niet weet of hij in de nieuwe beleidsperiode van de Nederlandse Taalunie (2013-2017) dezelfde toelage krijgt. “Wel zegt de Taalunie dat het samengaan met de KB geen negatieve effecten heeft op de continuïteit van de website en van onze expertise en dienstverlening”, verklaart hij.

ZO MIN MOGELIJK PAPIEREN TUSSENSTADIA

Niettemin zal de DBNL ooit de complete Nederlandse literatuur aanbieden. Met dat doel is de stichting Digitale Bibliotheek voor de Nederlandse Letteren in 1999 opgericht door de Maatschappij der Nederlandse Letterkunde, een jaar later ging de site online. Afhankelijk van de definitie van literatuur gaat het om twintig tot veertig miljoen gedigitaliseerde pagina's. Door elke maand nieuwe teksten toe te voegen, komt dat aantal vanzelf in zicht. De maandelijkse updates die de DBNL per mail verstuurt, tellen iedere keer ruim honderd nieuwe titels. Op dit moment gaat Klapwijk ervan uit zeker de tijd tot na 2020 nodig te hebben.

In totaal zijn nu tien miljoen pagina's digitaal beschikbaar, vertelt Klapwijk in het Leidse kantoor boven een Engelse pub dat de DBNL binnenkort verlaat. Dertig procent daarvan zijn gecontroleerde *xml*-pagina's, de rest zijn scans. Ook die zijn als *pdf*- of *txt*-bestand op woordniveau te doorzoeken, maar deze zijn aanzienlijk minder populair: ze worden vijf tot tien keer minder opgevraagd. “We scannen door, maar het tempo is nu wat gematigd”, zegt Van Stipriaan. “We hebben nu een goede werkvoorraad om uit te putten voor allerlei projecten.”

Het tempo om de met de *optical character recognition*-techniek gefabriceerde scans op te waarderen tot volwaardige *xml*-bestanden, voorzien van metadata, is wel omhoog gegaan. “Het digitaliseren gebeurt steeds efficiënter”, zegt Van Stipriaan. “Met almaar meer geavanceerde productiemethoden kunnen we steeds sneller steeds betere *xml* genereren. Door veel gelijkvormige massa aan te bieden, hebben we scherpe tarieven kunnen bedingen. De laatste slag die we nu willen maken is om een te digitaliseren boek bij wijze van spreken over de lopende band te laten lopen, via een zogenaamde *web based* productiemethode, met zo min mogelijk papieren tussenstadia. Daar gaat de

DBNL in 2013 mee experimenteren, maar dat vraagt wat van de organisatie – intern én bij de toeleveranciers.”

De mythe wil dat ieder boek met de hand wordt overgetikt in de Filippijnen. Dat is niet waar. Bedrijven “uit alle windstreken” – van Thailand en India tot Vietnam en de Filippijnen – scannen teksten met een computer. Alleen moeilijke boeken, zoals een in Gotische letter gedrukt liedboekje uit de zestiende eeuw, worden met de hand gedigitaliseerd. Mensen controleren de scans wel en bewerken die zo nodig redactioneel. De gepubliceerde tekst moet voor 99,995 procent foutloos zijn. “Dat is de hoogst verkrijgbare norm in de markt”, zegt Klapwijk. “Dat betekent: één fout per tien pagina’s. Wij monitoren alle bedrijven om te zien of ze aan de kwaliteitseis voldoen.”

Voor de keuze van de te digitaliseren teksten beschikt de DBNL over acht commissies die steeds nieuwe verlanglijstjes opstellen. Vier commissies behandelen aparte tijdvakken: de middeleeuwen, de gouden eeuw, achttiende/negentiende eeuw en twintigste/eenentwintigste eeuw. En vier commissies adviseren over specifieke genres: taalkunde, jeugdliteratuur, Surinaamse literatuur en Zuid-Afrikaanse literatuur. Prominente Nederlandse en Vlaamse neerlandici hebben zitting in een commissie: Nicoline van der Sijs (taalkunde), Anne Marie Musschoot (twintigste/eenentwintigste eeuw), Michiel van Kempen (Surinaamse literatuur) en vele anderen.

Of de voorkeuren van de commissies kunnen worden gehonoreerd, hangt af van de verkrijgbaarheid van boeken. “De eerste helft is makkelijk”, zegt Klapwijk, “het derde kwart gaat ook nog, maar daarna kan het lastig worden om iets in huis te krijgen. Gelukkig hebben we goede relaties opgebouwd met bibliotheken, erfgoedinstellingen en musea in Nederland en Vlaanderen. Aanvankelijk haalden we de meeste boeken bij de universiteitsbibliotheken, maar allengs kwamen meer partijen in beeld. De samenwerking met al deze partijen verloopt wonderwel goed. Omdat andere partijen de scans ook krijgen, is het goed mogelijk om redelijke deals te sluiten.”

Bijzondere samenwerkingsverbanden zijn die met partijen die zelf geld meebrengen om een bepaald segment van de Nederlandse literatuur te digitaliseren. Zeker als het past in het eigen beleidsplan van de DBNL zoekt Klapwijk actief naar zulke projecten. Voorbeelden zijn het digitaliseren van het complete oeuvre van Menno ter Braak in samenwerking met de Stichting Menno ter Braak. Of het beschikbaar maken van een kleine honderd klassiekers uit de Friese literatuur in samenwerking met het Fries historisch en letterkundig centrum Tresoar. Een plan om met het Fonds BKVB (nu Mondriaan Fonds) een canon van honderd strips aan te bieden, ging niet door.

“We zouden graag aan een aantal genres meer aandacht besteden om de spanwijdte van de site te verbreden”, zegt Van Stipriaan. “Gezien onze academische achtergrond ligt de nadruk op deftige literatuur, ook als het gaat om secundaire literatuur. De site bevat veel edities en studies. Het zou mooi zijn als we ook de brede laag daaronder

ruimte kunnen geven. Denk aan ondergeschoven kindjes zoals detectives. We blijven op dit punt bezig, met het project rond Limburgse literatuur waarvoor op 31 januari 2013 een afdeling binnen de DBNL is opgeleverd, en met een groot project met het Meertens Instituut en de Universiteit Utrecht rond liederen en liedjes. Welke stappen we precies kunnen zetten, is natuurlijk sterk afhankelijk van initiatieven uit het veld.”

ERNSTIG BESCHIMMELD, TOCH DIGITALISEREN

Ook de naamsbekendheid bij het grote publiek draagt bij aan het makkelijk verkrijgen van teksten. “Men heeft de neiging ons op te zoeken”, zegt Klapwijk. “Ook krijgen wij suggesties van individuele auteurs. Die voorstellen proberen wij te accommoderen als het goed past in ons programma.” Van Stipriaan vult aan: “In beginsel zullen wij wat wij krijgen aangeboden, proberen in te nemen. Alleen geen onafgemaakte manuscripten die wij als een uitgever persklaar moeten maken. Wat wij opnemen, moet ooit al door een uitgeverij interessant genoeg bevonden zijn om er een oplage van te maken.”

Desondanks krijgen de medewerkers van de DBNL niet vanzelf de complete Nederlandstalige literatuur in huis. In mei 2011 deden ze nog een oproep aan particulieren om oude boeken en tijdschriften, het liefst van voor 1940, op te sturen. Auteurs werd gevraagd hun eigen teksten in te sturen. Dat hebben ze geweten, vertelt Van Stipriaan. “Sommigen stuurden drie boeken op, anderen ruimden hun hele garage leeg, waar de tijdschriften – soms ernstig beschimmeld – hoog opgetast lagen. Bij elkaar waren het vrachtwagens vol. En van de Erfgoedbibliotheek Hendrik Conscience in Antwerpen kregen we al hun doublures, die ze wilden afstoten.”

Daarbij hoeft de DBNL niet per se alle teksten zelf te digitaliseren. Er zijn talloze instellingen die oude Nederlandstalige teksten scannen – dat werk moet niet dubbel worden gedaan. “Er was een tijd dat iedereen aan het digitaliseren sloeg”, zegt Klapwijk. “Dat gebeurde vooral om redenen van conservatie, de ontsluiting kwam op het tweede plan. Liedbundels bij de Universiteit Utrecht, kinderboeken bij de Stadsbibliotheek Haarlem, duizenden kinderboeken bij de Bibliotheek Rotterdam. En zo verder. Ook de Koninklijke Bibliotheek, die later een coördinerende rol kreeg in de massadigitalisering in Nederland, doet veel. Wij houden nauwgezet in de gaten wat er bij hen uit rolt.”

Ook de verhuizing van de DBNL naar KB moet in eerste instantie worden begrepen vanuit de wens om centraal één digitale bibliotheek te creëren. “Dat is zeker het perspectief van het Ministerie van Onderwijs, Cultuur en Wetenschap”, zegt Klapwijk. “Daarom gaan ook de organisaties van de openbare bibliotheek die bezig zijn met het opzetten van een landelijke bibliotheek voor het uitlenen van onder meer e-boeken, naar de KB. Het ligt voor de hand dat wij onze expertise voor het maken van betrouw-

bare tekstbestanden ook gaan inzetten om de kwaliteit van de *quick and dirty* gemaakte bestanden van de KB en van andere erfgoedinstellingen te vergroten. Dat bepleiten we al jaren, en als het nu echt gebeurt, zou dat voor de gebruiker een belangrijke stap voorwaarts zijn.”

De DBNL blijft wel een zelfstandige, mede door de Nederlandse Taalunie gedragen site, naast de talrijke projecten van de KB zoals de historische krantenbank en *Early Dutch Books Online*. Maar op den duur, zo verwacht Klapwijk, worden de collecties vindbaar gemaakt in één grote metasite, waar bijvoorbeeld ook de databank aan beeld- en geluidsfragmenten van het Instituut voor Beeld en Geluid te vinden zal zijn. “Bezoekers van ons zouden dan moeiteloos moeten kunnen doorklikken naar recensies van de werken, interviews met de auteur, verfilmingen, documentaires, enzovoort.”

Heel actueel is de DBNL niet. Uit de eenentwintigste eeuw, toch alweer voor bijna een achtste voorbij, is weinig beschikbaar. Hoofdzakelijk literaire tijdschriften als *Passionate*, *Tirade* en *Het Liegend Konijn* en een handvol non-fictietitels als de biografie van Menno ter Braak van Léon Hanssen en een pamflet van Geert Mak. Van hedendaagse auteurs is bijna niets te lezen. Van Arnon Grunberg: alleen tijdschriftbijdragen. Van A.F.Th. van der Heijden: wat korte teksten, waaronder een acht pagina's tellend fragment uit *De tandeloze tijd*. Zelfs van een oudere auteur als de onlangs overleden Bernlef, gedebuteerd in 1960, is niet één origineel werk te lezen.

De DBNL is voor het gebruik hiervan afhankelijk van de rechthebbenden. In principe werken auteurs en uitgeverijen graag samen. Slechts af en toe weigert een auteur of zijn nabestaande de opname van een al wat oudere tekst in de DBNL omdat hij denkt er nog aan te kunnen verdienen. In september 2011 sloot de DBNL de eerste overeenkomsten met uitgeverijen om hun backlist gezamenlijk te digitaliseren. Het gaat om WPG Uitgevers (waartoe gerenommeerde uitgeefhuizen als De Bezige Bij, Querido, Arbeiderspers en Standaard Uitgeverij behoren) en Lannoo-dochter Meulenhoff Boekeriej. Beide projecten zijn inmiddels afgerond.

Essentieel voor het succes van zo'n samenwerking is ten eerste een succesvol tegengaan van piraterij en dus inkomstenderving voor auteurs en uitgevers. In principe zijn nu nog alle titels, ook de auteursrechtelijk beschermde, vrij te bekijken en te downloaden. Het heeft volgens Klapwijk, die misbruik van DBNL-bestanden goed monitort, geleid tot niet meer dan een paar gevallen van illegaal verspreiden. Ook hebben recente horrorverhalen over de forse groei van digitaal misbruik niet tot gevolg dat auteurs minder animo tonen in het geven van toestemming of dat uitgevers gebruik maken van de *opt out*-mogelijkheid. Als dat al gebeurde, was dat omdat er een heruitgave kwam.

Het belangrijkste is echter dat de DBNL vanaf dit jaar een bijdrage levert aan de verspreiding van titels die in digitale vorm commercieel geëxploiteerd worden. Het is de bedoeling dat gebruikers van de honderden teksten die WPG en Meulenhoff Boekeriej hebben aangeleverd, deze volledig kunnen doorzoeken en fragmenten kunnen lezen, maar dat zij de volledige tekst alleen kunnen raadplegen tegen betaling. De DBNL linkt hen dan door naar een webwinkel of naar een voorziening voor digitaal lenen van de openbare bibliotheek. Van Stipriaan: “Het voordeel voor ons is dat wij het aanbod zo completer maken en ook dit type recent erfgoed beschikbaar kunnen maken ten behoeve van onderwijs en onderzoek.”

Met het project rond Friese literatuur heeft de DBNL twee jaar geleden voor het eerst ervaring opgedaan met het faciliteren van commerciële exploitatie. Door de geringe omvang van de Friese markt komen klassiekers als *It jubelijer* van Simke Kloosterman (1927) niet in aanmerking voor normale heruitgave. Het geïnteresseerde publiek kon ze daarom via de DBNL-site bestellen, waarna ze *on demand* werden geproduceerd door de Friese *print on demand*-uitgeverij Elikser. Later kwamen ze ook als e-boek te koop. Het project kent naar verhouding veel succes. “In negen maanden waren de eerste duizend boeken verkocht”, zegt Van Stipriaan. “Sindsdien loopt het aardig door.”

Dit model wordt ook toegepast voor Zuid-Afrikaanse literatuur, met het in Kaapstad gevestigde NB Publishers, en in de nabije toekomst met Limburgse literatuur. Hiervoor ontvangt de DBNL een *affiliate fee*. Een bescheiden fee, benadrukt Klapwijk, zodat alle marktpartijen hun normale marge kunnen krijgen. “Dit zal geen grote inkomstenstroom worden”, zegt hij. “Wij willen zelf geen boekhandel zijn, maar derde partijen faciliteren vanuit de overtuiging dat zowel zij als ons publiek daarbij gebaat zijn. Wij merken ook uit de post dat een belangrijk deel van de vragen van het publiek gaat over: waar kan ik dit of dat boek krijgen?”

TOT 20.000 BEZOEKERS PER DAG

De bezoekcijfers van de DBNL bewijzen dat zij commerciële partijen een groot publiek kunnen bieden. Vanaf het begin in 2000 groeide het onophoudelijk. “De eerste jaren verdubbelde het gebruik ieder jaar”, zegt Van Stipriaan. “Ook nadat rond 2004 de hele wereld ging digitaliseren, Google voorop, hield de groei aan, maar die vlakke wel wat af naar tien procent op jaarbasis. Nadat we een nieuwe versie van de site hadden gelanceerd in 2010 en het bezoek in onze statistieken waarschijnlijk anders berekend werd, daalden de bezoekcijfers opeens veertig procent. Toen we een jaar later voor het eerst weer goed konden vergelijken, bleken we gelukkig weer tien procent te zijn gegroeid. En het laatste jaar is het bezoek weer enorm gestegen: plus vijftieng procent.”

Concreet komt dat neer op: 15.000 tot 20.000 bezoekers per dag, waarvan een onevenredig groot aantal uit Vlaanderen, die relatief lang op de site blijven. Met gemiddeld tien bekeken pagina's per bezoeker, komt dat neer op 150.000 tot 200.000 bekeken pagina's per dag. Het totale aantal bekeken pagina's in 2012 bedroeg zestig miljoen. "Het is alleen wat sinister dat de groei parallel lijkt te lopen met het verval van het boekenbedrijf", zegt Van Stipriaan. "Want ik denk dat de opkomst van de tablet-computer een belangrijke verklaring voor de recente groeispurt is. Dankzij de iPads en andere toestellen kun je voor het eerst gemakkelijk in je stoel een tekst lezen. Helaas kunnen we de verklaring niet heel goed bestuderen."

De DBNL zou dan ook maar wat graag een gebruikersonderzoek uitvoeren. Wie zijn de gebruikers en wat zijn hun behoeftes waaraan de DBNL misschien beter kan voldoen? "We weten wel op welke zoektermen bezoekers binnenkomen", zegt Van Stipriaan. "Bij termen als 'Dikkertje Dap', 'Nederlandse spreekwoorden' en vreemd genoeg 'hulpwerkwoorden' staan wij bovenaan bij Google. Maar de bezoekers zelf zijn anoniem. Ja, een groot deel zijn scholieren, maar hoeveel? En met welk motief komen ze bij ons? Hoe wordt de site dan in het onderwijs gebruikt? Door welke scholen? Helaas is een goed gebruikersonderzoek buitengewoon kostbaar."

Toch is dat niet de eerste prioriteit. Dat blijft onverminderd het opbouwen van een complete databank van de Nederlandse literatuur vanaf *Hebban olla vogala* – met ingang van dit jaar met extra nadruk op het vergroten van het aandeel Vlaamse literatuur in de DBNL. Ooit komen de makers er toch wel achter wie *Monsieur Hawarden* leest voor zijn plezier of *Mensen achter de dijk* bestudeert voor wetenschappelijk onderzoek. Informatie die dan zal worden gebruikt voor nieuwe functionaliteiten, om de geïnteresseerden in Filip De Pillecyn nog beter van dienst te zijn.

www.dbnl.org
