

Erfgoed beschikbaar stellen

Open data in de Koninklijke Bibliotheek

Als nationale bibliotheek verzamelt de Koninklijke Bibliotheek (KB) alle publicaties van en over Nederland, van handschriften uit de Middeleeuwen tot de krant van gisteren. Het is onze taak om dit erfgoed zo breed en open mogelijk beschikbaar, vindbaar en herbruikbaar te maken. Dit artikel geeft in vogelvlucht weer hoe de KB dit aanpakt.

TEKST OLAF JANSSEN

Digitale collectie



De KB is sinds 1798 de nationale bibliotheek van Nederland. We verzamelen alle publicaties uit en over Nederland, van manuscripten uit de 11e eeuw tot hedendaagse kranten. Dit gebeurt voor een groot deel nog steeds fysiek, maar in toenemende mate worden ook digitale publicaties aan de collectie van de KB toegevoegd.

De uitbreiding van onze digitale collectie gebeurt enerzijds door *digital born*-materiaal duurzaam te bewaren, denk aan e-books of digitale wetenschappelijke tijdschriften. Anderzijds zetten we vol in op het digitaliseren van onze fysieke publicaties. Zo worden binnen Metamorfoze, het nationale programma voor het behoud van het papieren erfgoed¹, duizenden historische boeken, kranten en tijdschriften *full-text* gedigitaliseerd. Bovendien hebben we sinds 2010 een overeenkomst met Google, waardoor tien-

duizenden Nederlandse boeken van voor 1900 in bits en bytes worden omgezet.²

Online ondanks auteursrechten



Gezien onze publieke taakstelling stellen we deze gedigitaliseerde publicaties zoveel mogelijk online beschikbaar. We doen dit hoofdzakelijk via Delfher, de landelijke website voor full-text gedigitaliseerde historische Nederlandse kranten, boeken en tijdschriften uit bibliotheken, musea en andere erfgoedinstellingen.³

We zijn daarbij gebonden aan wat de Auteurswet ons toestaat. Met name op het 20e-eeuwse deel van onze collectie rusten vaak nog rechten, waardoor we die publicaties niet zomaar op Delfher mogen zetten. De KB is in voortdurende onderhandeling met rechthebbenden, zoals schrijvers, fotografen en uitgevers, om collectieve licentieovereenkomsten voor een bepaalde duur af te sluiten. En vaak met succes, hierdoor zijn bijvoorbeeld kranten uit de jaren negentig op dit moment voor iedereen gratis in Delfher te zien.

Publiek domein



Op publicaties waarvan de makers langer dan 70 jaar dood zijn, of die meer dan 70 jaar geleden zijn uitgegeven, rusten geen auteursrechten meer. Ze behoren toe aan het publieke domein. Daardoor mag iedereen ermee doen wat hij wil. Dit materiaal mag dus zonder beperkingen of voorwaarden als open data online gezet, gedeeld, bewerkt, heruitgegeven en verkocht worden.

Voor de KB is dit bijzonder gunstig, omdat we als erfgoedinstelling veel historisch materiaal bewaren. Niet alleen kranten, boeken en tijdschriften van vóór (plusminus) 1900, maar ook bijvoorbeeld middeleeuwse handschriften, atlanten uit de Gouden Eeuw of de Handelingen van de Staten-Generaal sinds 1814.

Open data via eigen diensten

Dit rechtenvrije materiaal mogen we dus ongehinderd digitaliseren en via onze eigen webdiensten aanbieden. Dit doen we dan ook volop. Zo behoort een aanzienlijk deel van de publicaties in Delpher tot het publieke domein. Bij elke rechtenvrije publicatie geven we manifest in woord (“Er rusten geen rechten meer op dit object. Daarom mag het door iedereen vrij worden (her)gebruikt voor ieder doel.”) en in beeld (door middel van een publiek domein-icoon) aan dat het vrijelijk hergebruikt mag worden.

De KB biedt rechtenvrij materiaal aan via de eigen webdiensten

De KB wordt regelmatig benaderd door onder anderen taalkundigen of datawetenschappers die grote hoeveelheden publicaties willen analyseren, denk aan teksten van tienduizenden krantenafleveringen. Zij willen die teksten zonder al te veel technisch gedoe in één keer kunnen downloaden. Speciaal voor dit soort niet-technische bulkhergebruikers hebben we het Delpher Open Krantenarchief in het leven geroepen. Hier kunnen de teksten van alle rechtenvrije kranten (van vóór 1877) eenvoudig als open datasets (zip-bestanden) gedownload worden.⁴

Een ander deel van onze open data bieden we aan via open download- en zoek-API's. Enigszins technisch onderlegde gebruikers kunnen via deze kanalen bijv. duizenden rechtenvrije boeken uit de 18e eeuw downloaden.⁵

Ten slotte bieden we via data.bibliotheken.nl een deel van onze bibliografische databases en thesauri als linked open data aan. Hiermee sluiten we aan bij de ontwikkelingen rondom het semantische web.

Externe platformen en netwerken



Het aanbieden van open data via onze eigen diensten is slechts een beginpunt, omdat die in de praktijk een beperkt bereik blijken te hebben. Als publieke instelling willen we zoveel mogelijk mensen bereiken. We sluiten daarom ook aan bij bredere platformen en netwerken, niet alleen bij die van de erfgoedsector en de overheid, maar ook bij die van grote (inter)nationale gemeenschappen. Zo maken we onze data onder andere beschikbaar via

• *Open Cultuur Data*, een netwerk dat zich inzet voor het openstellen van open data uit de culturele sector en de ontwikkeling van nieuwe waardevolle toepassingen daarmee.⁶

• *Netwerk Oorlogsbronnen*, een samenwerkingsverband van ruim 430 instellingen om de vindbaarheid en bruikbaarheid van de Nederlandse collectie over de Tweede Wereldoorlog te verbeteren.⁷

• *data.overheid.nl*, een register met informatie over en verwijzingen naar datasets van Nederlandse overheden.

• *Europeana*, het Europese platform en netwerk voor digitaal cultureel erfgoed.⁸



Hergebruik stimuleren, maatschappelijke meerwaarde

We vinden het als KB erg belangrijk dat anderen kunnen voortbouwen op onze open data, we zien het met andere woorden als een halffabricaat en geen doel op zich. Het passief aanbieden van data via eigen en externe diensten is vaak niet genoeg om de gewenste maatschappelijke meerwaarde te krijgen. Daarom stimuleren we het hergebruik van onze rechtenvrije content actief.

Als wetenschappelijke instelling onderhoudt de KB nauwe banden met universiteiten en onderzoeksprojecten, met name in de wereld van de geesteswetenschappen. Vooral binnen de opkomende tak van de *digital humanities*⁹ profileren we ons als leverancier van open data. Zo zijn binnen het project Polimedia onze data gebruikt om in beeld te brengen hoe er tussen 1945 en 1995 in kranten en op de radio verslag werd gedaan van de debatten in de Tweede Kamer.¹⁰ En in het internationale onderzoeksproject Oceanic Exchanges wordt met computertechnieken onderzocht hoe nieuws uit de 19e eeuw zich mondiaal verspreidde via kranten. Hierbij worden onder andere de kranten uit het Delpher Open Krantenarchief gebruikt.¹¹

Actief hergebruik van de eigen rechtenvrije content wordt gestimuleerd

We brengen onze datasets ook actief onder de aandacht bij programmeurs, appbouwers, designers en data-analisten. Speciaal voor deze doelgroepen organiseren we regelmatig *hackathons*, evenementen waarbij we dit type gebruikers uitnodigen om binnen een afgebakende tijd innovatieve ideeën en/of software te ontwikkelen op basis van (open) data uit erfgoedcollecties. Eerder dit jaar organiseerden we samen met het Rijksmuseum HackaLOD 2018, een 24 uren-hackathon rondom (linked) open data van culturele instellingen.¹²

Net als andere publieke erfgoed- en kennisinstellingen bevest de KB steeds beter hoe belangrijk het publieke domein is voor de onze maatschappelijke taakuitvoering. We proberen dan ook zoveel mogelijk andere partijen bewust(er) te maken van de meerwaarde die open data voor hen kunnen hebben. Daarom was de KB in januari partner van de eerste Nederlandse Publiekdomeindag. Tijdens dit geslaagde evenement werd stilgestaan bij het publieke domein en de makers van wie de werken op 1 januari dit jaar auteursrechtvrij werden.¹³

Wikipedia



Een van de meest effectieve manieren om meer mensen makkelijker in aanraking te laten komen met Nederlandse geschiedenis, cultuur, wetenschap en andere onderwerpen in onze collectie is via Wikipedia. Deze encyclopedie is veruit de meest zichtbare verzameling feitelijke informatie op het internet en wordt door ruim 80% van de Nederlanders gebruikt. Het gemiddelde Wikipedia-artikel is ook bijzonder goed via Google en andere zoekmachines te vinden, het staat vaak bij de eerste vijf zoekresultaten.

clopedie is veruit de meest zichtbare verzameling feitelijke informatie op het internet en wordt door ruim 80% van de Nederlanders gebruikt. Het gemiddelde Wikipedia-artikel is ook bijzonder goed via Google en andere zoekmachines te vinden, het staat vaak bij de eerste vijf zoekresultaten.

Wikipedia past ideaal binnen de open-datastrategie van de KB

Sinds 2013 werkt de KB nauw samen met de mensen die Wikipedia maken, veelal vrijwilligers in binnen- en buitenland die er in hun eigen tijd aan werken. Deze internationale Wikipedia-gemeenschap gebruikt bijvoorbeeld rechtenvrije afbeeldingen en teksten uit de KB-collectie om de kwaliteit van de encyclopedie te verbeteren. En daar profiteren weer miljoenen mensen van! Bovendien is Wikipedia op zijn beurt een platform voor open data, want alle teksten, beelden en data in de encyclopedie – ook die van de KB – mogen door iedereen verder hergebruikt en verspreid worden.¹⁴ Wikipedia past daardoor ideaal binnen de opendatastrategie van de KB. ●



Olaf Janssen
Olaf.Janssen@KB.nl

Olaf Janssen is als Open Data & Wikimedia-coördinator werkzaam bij de Koninklijke Bibliotheek in Den Haag.¹⁵

Noten

- 1 Metamorfoze, www.metamorfoze.nl
- 2 Samenwerking KB en Google, www.kb.nl/nieuws/2017/kb-en-googele-doen-samen-nog-een-digitaliseringsproject
- 3 Delpher, www.delpher.nl
- 4 Data in Delpher & Delpher Open Krantenarchief, www.delpher.nl/data
- 5 KB Dataservices & open APIs, www.kb.nl/dataservices
- 6 Open Cultuur Data, www.opencultuurdata.nl
- 7 Netwerk Oorlogsbronnen, www.oorlogsbronnen.nl
- 8 Europeana, www.europeana.eu
- 9 Digital humanities, en.wikipedia.org/wiki/Digital_humanities
- 10 Polimedia, www.polimedia.nl
- 11 Oceanic Exchanges, diggingintodata.org/awards/2016/project-oceanic-exchanges-tracing-global-information-networks-historical-newspaper
- 12 HackaLOD, www.hackalod.com
- 13 Publiekdomeindag 2018, www.publiekdomeindag.nl
- 14 Samenwerking KB met Wikipedia, www.kb.nl/wikipedia
- 15  De auteur stelt dit artikel beschikbaar onder de licentie Creative Commons Naamsvermelding-GelijkDelen 4.0. Zie <https://creativecommons.org/licenses/by-sa/4.0/deed.nl> voor aanvullende informatie.

Zie ook www.od-online.nl