

Pagina 3 • Oude publicaties CBS

Het CBS is bezig met een project om oude publicaties te digitaliseren en via internet toegankelijk te maken. Een half miljoen pagina's staan op het programma.

Pagina 4 • DANS bouwt resolver

Een resolver verbindt een blijvende identificatiecode met de plek op het internet waar de inhoud achter die code zich bevindt. DANS heeft zo'n resolver gebouwd.



Pagina 5 • Franciska de Jong

De Twentse hoogleraar taaltechnologie Franciska de Jong combineert taalkunde en statistiek bij het werken aan multimediazoektechnologie, bijvoorbeeld om in audio-interviews te kunnen zoeken.

Pagina 6 • Vergrijzingsonderzoek

Netspar richt zich op de sociale en economische vraagstukken rond een vergrijzende bevolking. Door middel van onderzoek, onderwijs en kennisuitwisseling.

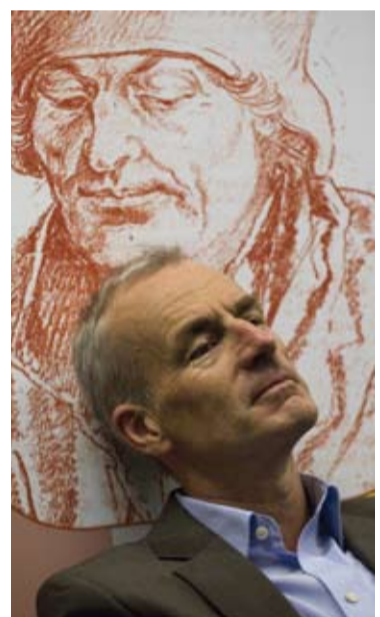
EN VERDER

Agenda.....	2
Nieuws.....	3
Achtergrond.....	4
Ontwikkelingen.....	7
Column.....	8
Gelezen.....	8

Spinoff van nieuwe BNTL: samengebruik van bibliografieën

Op 24 april is in Den Haag de nieuwe Bibliografie van de Nederlandse Taal- en Literatuurwetenschap (BNTL) gelanceerd. De 'officiële handeling' werd verricht door Karina van Dalen, verantwoordelijk bij de Nederlandse host van de bibliografie, het Huygens Instituut, en directeur Marcus de Schepper van de Belgische partner, het Bureau van de Bibliografie van de Neerlandistiek.

De nieuwe BNTL is de opvolger van de versie die sinds 1970 werd opgebouwd vanuit het Huygens Instituut en de Vlaamse Bibliografie voor de Neerlandistiek. Hij bevat Nederlandse, Vlaamse en buiten-



Een tevreden Huygens-directeur Henk Wals

landse titels van publicaties over Nederlandse taal- en letterkunde uit de periode 1940 tot heden (dus geen primaire literatuur), waar mogelijk met een link naar de digitale volledige teksten.

Die BNTL is sinds de jaren tachtig van de vorige eeuw digitaal maar daarmee nog niet gemakkelijk toegankelijk, legt Van Dalen uit. De bibliografie werd toen ondergebracht in het bibliotheekstelsel Pica van OCLC, 'maar daar konden alleen abonneerders bij. De versie die sinds vorig najaar intern draait en die nu officieel is gelanceerd, is voor iedereen toegankelijk en dat is een groot verschil. Eén van de aspecten waar wij ons op verheugen is dat vanaf het komend najaar ook gebruikers gegevens kunnen invoeren'. Daarmee wordt tegelijk een grotere inhoudelijke dekking bereikt en iets gecompenseerd van het verminderde budget voor de BNTL.



De Schepper en Van Dalen verrichten de opening

Van Dalen: 'Het systeem is in deze nieuwe versie vereenvoudigd. Er zijn minder velden om in te vullen. Daardoor kunnen we nu veel sneller titels invoeren.' Voor de mogelijkheid om extern in te voeren wordt een flinke belangstelling verwacht. Die mogelijkheid heeft ook de vorm van een product aangenomen: de Collaborative Online Bibliography (COB), die al operationeel is voor de Friese Bibliografie en waarvan een Engelse versie wordt onderzocht voor de naamkundebibliografie van de International Council of Onomastic sciences (ICOS).

De nieuwe BNTL is in ontwikkeling, benadrukken de makers. De bedoeling is om regelmatig mogelijkheden aan het systeem toe te voegen. Op dit moment kan gelijktijdig worden gezocht op informatie uit de titels (auteur, jaar, titel) en toegevoegde informatie (trefwoorden, samenvattingen en andere informatie). Ook kan er geavanceerd worden gezocht op specifieke onderdelen. De digitale teksten waarnaar wordt verwezen zullen later dit jaar ook doorzoekbaar zijn. Dagelijks worden titels en teksten door de interne redactie toegevoegd. (MdG)

Data uit beleidsonderzoek voor rijk: verplicht archiveren

Gegevens uit beleidsgericht onderzoek in opdracht van de rijksoverheid, 'die zich lenen voor meervoudig gebruik', moeten voortaan worden gearchiveerd bij DANS (Data Archiving and Networked Services). Dat staat in een modelovereenkomst die is verbonden aan de nieuwe Algemene Rijksvoorwaarden voor het verstrekken van Opdrachten tot het verrichten van Diensten (ARVODI).

De nieuwe voorwaarden zijn op 16 maart in de Staatscourant gepubliceerd. Ze leggen vast hoe er met de rijksoverheid zaken gedaan kunnen worden op alle gebieden waar die overheid gebruik maakt van diensten door derden. Een van die gebieden is het uitvoeren van onderzoek voor het beleid. Daarvoor is een aparte modelovereenkomst gemaakt. Die vraagt van de opdrachtnemers naast deponeren van de data bij DANS dat ze tevoren hun onderzoek aangeven bij de Nederlandse Onderzoeksdatabank NOD, die wordt onderhouden door de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW). Ook wordt het deponeren van het onderzoeksrapport bij het e-Depot van de Koninklijke Bibliotheek verplicht gesteld. Voor het deponeren van data maakt de modelovereenkomst

wel het voorbehoud dat die moeten zijn 'geschoond van vertrouwelijke gegevens, gerubriceerde gegevens en gegevens waarvan de publicatie op grond van de wet- en regelgeving achterwege moet blijven'. Data die door die zeef komen moeten 'binnen drie maanden na publicatie van het eindrapport door de opdrachtnemer (...) aan Data Archiving and Networked Services (DANS) worden aangeboden', aldus de modelovereenkomst. De Algemene Rijksvoorwaarden zijn opgesteld onder verantwoordelijkheid van het ministerie van Algemene Zaken. De uitwerking van de voorwaarden voor de leverantie van beleidsgericht onderzoek is gebeurd door het departement van Onderwijs, Cultuur en Wetenschap (OCW). Dat verzekert desgevraagd dat zowel de voorwaarden als de daaraan verbonden

modelovereenkomsten bedoeld zijn voor gebruik door alle departementen. DANS is nog in gesprek met het ministerie van OCW over de praktische uitwerking van het verplicht deponeren. (MdG)



Minister Plasterk, verantwoordelijk voor de voorwaarden voor beleidsgericht onderzoek

Microsoft add-in beste conversietool PDF/a

De Microsoft add-in geeft de beste resultaten bij het converteren van MS Office formaten naar het veel duurzamere PDF/a formaat. Dat blijkt uit onderzoek van de Koninklijke Bibliotheek. PdfMaker van Adobe is een goede tweede. Voor wie minder hecht aan de exacte weergave van het originele uiterlijk maar meer aan structuur, bleek conversie via OpenOffice een goed alternatief. Tools die een printerdriver gebruiken om een PDF te converteren, zijn geschikt voor een natuurgetrouwe weergave van het uiterlijk en platte tekst, maar niet voor gebruikers die de structuur van het document belangrijk vinden. Overigens was geen van de geteste tools in staat om een testbestand te

converteren tot een valide PDF/a, wat betekent dat geen van de conversieresultaten voldeed aan de officiële PDF/a-grammatica.

De Koninklijke Bibliotheek besloot om een aantal conversietools te testen omdat het bestandsformaat waarin wordt opgeslagen grote invloed heeft op de houdbaarheid van digitale informatie. Eerder onderzoek had al aangetoond dat het PDF-formaat voor archiveren, PDF/a, in die zin veel betere toekomstperspectieven biedt dan de MS Office formaten. Bij deze test werden de resultaatbestanden beoordeeld op de drie eigenschappen *inhoud*, *uiterlijk* en *structuur*.

Informatie: wouter.kool@kb.nl. (IA)