Data uit psychologie: bewaren is geen gebruik maar zou het wel moeten zijn

Hoe gaan academische psychologen om met hun onderzoeksdata? De gangbare indruk is dat ze zich weinig aan archieven gelegen laten liggen. De Amsterdamse psycholoog Jelte Wicherts bevestigde die indruk onlangs toen uit zijn onderzoek bleek dat 'psychologen niet snel bereid zijn om anderen toegang te verlenen tot hun databestanden'. Hij vroeg om de bestanden bij 141 gepubliceerde artikelen en stuitte op een non-response van bijna driekwart. Jacques Dane, hoofd van het Groninger Archief- en Documentatiecentrum Nederlandse psychologie (ADNP) spreekt over 'het geheugenverlies' van de psychologie. 'Geen traditie', analyseerde hij eerder in *e-data&research*.

DANS, het nationale data-instituut, wilde meer hebben dan een indruk en entameerde een onderzoek. 'We willen weten in hoeverre academische psychologen hun gegevens opslaan of archiveren, dat wil zeggen ergens onderbrengen met metadata zodat ze voor anderen terug te vinden zijn', zegt onderzoekster Caroline Voorbrood. 'En we zijn benieuwd of men behoefte heeft aan het delen van data; ook de eigen data'. Onder leiding van themamanager Heleen van Luijn, net als zijzelf psychologe, begon ze haar onderzoek in de loop van oktober. Intussen is ze al een eind op streek met een reeks van twintig interviews met hoogleraren en universitaire hoofddocenten aan verschillende universiteiten. Psychologen die in de markt opereren, bijvoorbeeld arbeids- en organisatiepsychologen binnen een van de vele selectiebureaus die ons land rijk is, worden niet in het onderzoek betrokken. Voorbrood: 'De academische onderzoekwereld is al groot genoeg, en bovendien is DANS nu eenmaal vooral op die wereld gericht'.

DAGMAR STIEBRAL

Caroline Voorbrood (l) en Heleen van Luijn

Enthousiasme

De *semi-gestructureerde interviews* vormen de ene helft van het onderzoek. Een schriftelijke enquête onder medewerkers van dezelfde onderzoekseenheden waar de gesprekken plaatsvonden is de andere helft. Mogelijk zal die enquête, waarvoor enkele honderden psychologen worden benaderd, een steviger basis kunnen leggen onder de indrukken die uit de serie interviews oprijzen. Welke die indrukken zijn willen Voorbrood en Van Luijn nog niet erg gedetailleerd zeggen: 'Het onderzoekverslag komt in de zomer'. Duidelijk is wel dat archiveren onder psychologen inderdaad geen dagelijks werk is, dat dat volgens veel onderzoekers eigenlijk anders zou moeten, en dat in bepaalde subdisciplines een grote bereidheid heerst om er ook wat aan te doen. 'De meeste onderzoekers erkennen wel dat je de data uit wetenschappelijk oogpunt beschikbaar moet houden', ziet Voorbrood, 'maar het gebeurt weinig, misschien ook omdat zulk onderzoek zelden wordt gedaan'. Data archiveren voor hergebruik in ander onderzoek wordt in sommige subdisciplines wel degelijk als heel nodig gezien. 'We merken verschillen tussen de subdisciplines. Die hebben te maken met het soort onderzoek', aldus de onderzoekster. 'Experimentele psychologen doen vaak studies met een klein aantal proefpersonen. Hun belang is heel anders dan bijvoorbeeld dat van degenen die met methoden en technieken bezig zijn. Daar merk je wel dat ze staan te trappelen om over alle data te kunnen beschikken die er maar zijn'. Voorbrood bespeurt in deze kring nu al een zeker enthousiasme en is dus niet bang dat haar onderzoekrapport, waarover ze ook graag in de vakpers wil publiceren, in een vacuüm terecht zal komen. (MdG)

Informatie: caroline.voorbrood@dans.knaw.nl of heleen.van.luijn@dans.knaw.nl

InterEdition: een begin van samenwerking rond digitale edities

Op initiatief van het Huygens Instituut vond eind januari in Den Haag een bijeenkomst plaats van specialisten op het gebied van digitale edities. Doel was het verkennen van de mogelijkheden om samen te werken.

Het maken van digitale edities van letterkundige en historische werken is niet eenvoudig. Hoewel er de afgelopen twintig jaar hard gewerkt is aan codeerstandaards, programmatuur voor tekstvergelijking, annotatiehulpmiddelen, publicatietools en analysesoftware is het maken van digitale edities nog niet voor de gemiddelde onderzoeker weggelegd. In de praktijk bestaat vaak de behoefte aan ondersteunende staf of speciale training die afleidt van het eigenlijke wetenschappelijke werk.

Alle betrokkenen weten dat alleen samenwerking aan deze situatie een eind kan maken. Digitale edities zijn niet interessant voor de Microsofts van deze wereld. Het zijn instellingen met bescheidener budgetten die aan edities werken, en taakverdeling is dus noodzakelijk (ook al kan soms worden meegelift op grootschaliger initiatieven voor digitale bibliotheken of het GRID). Op uitnodiging van het Huygens Instituut kwam 25 en 26 januari in Den Haag een aantal internationale specialisten bijeen voor een initiatief om te proberen samenwerking van de grond te tillen. Aanwezig waren vertegenwoordigers van het Amerikaanse Perseus, de grote digitale bibliotheek gespecialiseerd in de klassieken, uit het Verenigd Koninkrijk onder meer Peter Robinson, bekend van zijn werk aan de Canterbury Tales en de programma's Collate (tekstvergelijking) en Anastasia (publicatie van XML-bestanden), met zijn collega Barbara Bordalejo; uit Vlaanderen onder andere Edward Vanhoutte van het Centrum voor Teksteditie en Bronnenstudie, bekend van de digitale Streuvels-editie en het Digital Archive of Letters in Flanders en Dirk van Hulle, maker van een digitale Beckett-editie. Een van de deelnemers uit Duitsland was Fotis Jannidis, als vertegenwoordiger van het TextGrid-initiatief. Het Huygens Instituut zelf werkt aan een omgeving die de onderzoeker/tekstbezorger moet ondersteunen tijdens het hele traject van transcriptie van de bron tot analyse van het gepubliceerde materiaal. Uit Nederland was daarnaast ook DANS vertegenwoordigd.

Het TextGrid-initiatief is een interessante partner omdat het op kleinere schaal al een architectuur voor samenwerking inricht. Het bestaat namelijk uit een aantal samenwerkende universiteiten en andere betrokkenen. De verschillende componenten van het systeem zullen elkaar via web services aanroepen. In Den Haag is bijvoorbeeld gesproken over de componenten waaruit zo'n systeem zou moeten bestaan, over de communicatie tussen die componenten, over beveiliging, over de noodzaak van unieke *identifiers* voor teksten, over annotaties en over tekstvergelijking. Ook meer organisatorische aspecten kwamen aan de orde, zoals een verdeling van taken, potentiële partners en mogelijkheden voor financiering. Afgesproken is om onder de voorlopige naam 'InterEdition' te werken aan het afstemmen van de plannen en het vinden van een breder (inter-) nationaal draagvlak. (PB)

VGI innovatieprijs 2007

Ieder jaar reikt de Vereniging voor Geschiedenis en Informatica de VGI innovatieprijs uit voor de meest vernieuwende Vlaamse of Nederlandse ICT-toepassing op het terrein van geschiedenis en cultuur. In september 2006 werd de prijs nog uitgereikt aan het project Teleblik dat historische radio en televisie voor het onderwijs beschikbaar maakt. Ook in 2007 wil de VGI weer een innovatieprijs toekennen. Daarom doet de vereniging een oproep om innovatieve websites, applicaties of projecten aan te melden. Een deskundige jury zal drie Vlaamse of Nederlandse toepassingen nomineren die tot de meest vernieuwende van het jaar gerekend kunnen worden. Deze drie zullen op een speciale studiedag van de VGI in september worden gepresenteerd, waarna een winnaar gekozen zal worden. Voor meer informatie over de studiedagen en de VGI-innovatieprijs, aanmelding en vorige winnaars; zie www.vgi-online.org. (LS)

MIXED van DANS

(vervolg van pagina 3)

Toch zijn er weinig vergelijkbare projecten bekend die betrekking hebben op databases. Roorda: 'Het federale archief van Zwitserland is bezig geweest met het SIARD-project. Dat had een soortgelijke doelstelling, maar anders dan MIXED is het geen Open Source'.

Als MIXED tot het gewenste resultaat leidt zal het dus op grote belangstelling kunnen rekenen, bevestigt Roorda. Hij ziet dan ook naast een algemeen belang voor het archiveren van data een bijzonder belang voor DANS: 'Het archiefsysteem DANS EASY staat sinds kort open voor alle onderzoekers in de alfa- en gammawetenschappen.

Het is aantrekkelijk om daarin je gegevens onder te brengen. maar het zal nog veel aantrekkelijker worden als daar de programmatuur en procedures van MIXED aan worden toegevoegd. Dan krijgt de garantie van DANS concreet gestalte dat data die via EASY worden gearchiveerd ook duurzaam zijn opgeslagen'.

In de twee jaar die voor het project staan zit ook de zogenaamde nazorg: niet alleen moeten programmatuur en procedures dan gebruiksklaar zijn, maar ook zal er bekendheid aan worden gegeven in de wetenschappelijke wereld. MIXED heeft een budget van bijna twee miljoen euro, waarvan de kleinste helft wordt betaald door DANS zelf en de grootste via SenterNovem door het overheidsprogramma PRIMA (zie www.senternovem.nl). (MdG)