

Digiscepticus Peter van Emde Boas

'Er zijn altijd Farao's die Jozef niet meer gekend hebben'



Bij revoluties horen sceptici. Hoe groter de euforie over de nieuwe mogelijkheden, des te meer voelen onafhankelijke denkers zich soms aangespoord om te wijzen op de risico's die daarachter schuilgaan. Peter van Emde Boas is zo'n scepticus. Een digiscepticus.

Van Emde Boas is als mathematisch informaticus verbonden aan het Amsterdamse Institute for Logic, Language and Computation. De digitale wereld heeft voor hem niet veel geheimen en het is dan ook niet daartegen dat zijn bedenkingen zich richten. Daarvan getuigen ook de vele artefacten van digitale hardware die hij in zijn huis heeft verzameld; van een NewBrain handcomputer uit de jaren zeventig (zelf te programmeren) tot een IBM-pc met 286-processor die in de jaren tachtig nog als de top werd beschouwd op het gebied van rekenkracht voor onderzoekers.

Maar het zijn juist die machines die van Emde Boas nu gebruikt om zijn scepsis te illustreren. 'Laten we het maar eens proberen' zegt hij tegen zijn vrouw – ook wiskundige –

terwijl hij een floppy met gegevens van meer recente datum in de a: poort van een tien jaren oude Aptiva steekt. 'Kijken wat er gebeurt'. Het resultaat is te voorspellen. De voormalige topper geeft na minuten vruchteloos geratel te kennen zich geen raad te weten met de inhoud van de floppy. Van Emde Boas: 'In veel gevallen zijn files na enkele jaren al niet meer te zien of te lezen. Je bent afhankelijk van de lees- of afspelapparatuur. Volledige backward compatibiliteit wordt maar al te vaak over het hoofd gezien'. Hij noemt ook het voorbeeld van zijn eigen redelijk moderne Macintosh computer: niet meer in staat om gegevens te lezen die werden opgeslagen op diskette door haar eigen naamgenoot van vijfjaar geleden. Van Emde Boas: 'Er zijn steeds weer Farao's die Jozef niet meer kennen'.

Tegenwerpen dat papier ook kan verbranden en dat het dan vaak om unieke exemplaren gaat, dat videotapes tot stof kunnen vergaan en dat het vermenigvuldigen van zulke dragers veel omslachtiger is dan van

digitale bestanden, maken op de informaticus weinig indruk. 'Nog onlangs bleken we in staat na 1800 jaar het Judasevangelie te lezen'. Hij somt nog eens een rijtje fatale fouten op die recentelijk in het nieuws waren, zoals die van de veertig duizend belastingaangiften die met een druk op de knop verdwenen. 'De mens is tot het kwade geneigd', constateert hij met een grijns en daarom is het onvermijdelijk dat dit soort fouten wordt gemaakt.

'Als het gaat om grote bestanden en databanken ben je volledig afhankelijk van de wil en de intenties van de beheerders'. In het geval van de wetenschappelijke tijdschriften zijn dat commerciële bedrijven met dito doelstellingen, benadrukt de onderzoeker. 'En wie garandeert dat deze bedrijven over vijftig jaar nog bestaan en deze service willen verlenen?'

Maar een instelling als het zojuist opgerichte DANS, is die dan niet juist nuttig om precies deze bezwaren het hoofd te bieden? Het bewaren en voor de lange termijn veiligstellen van gegevensbestanden, daar gaat het DANS immers om... 'Het punt is niet dat ik daartegen ben', verzekert de digiscepticus. 'Het punt is dat je, om een dergelijke opdracht met succes uit te voeren, over een organisatiegraad moet beschikken die volgens mij niet binnen ons bereik ligt'.

Zou Van Emde Boas dan zijn eigen gegevens aan DANS ter beschikking stellen als hem daarom werd gevraagd? 'Ik zou op zo'n verzoek wel ja zeggen, maar ik zou er geen reden in zien om zelf mijn data niet meer op een toekomstbestendige wijze te bewaren'. (MdG)

Column

Ton van Rietbergen

DANS om de waarheid



April was de maand van de filosofie. Na dierendag, de dag van de secretaresse, de homodag en het weekend van de korenwolf (hoe gaat het trouwens met ze?) maakten de filosofen zich een maand lang druk over de waarheid. Na deze maand kunnen we gewoon weer een loopje met de waarheid nemen, maar in april moeten we de waarheid eren en bloemen voor haar meenemen. In het licht van de aanwezigheid van de eerste april is de maand wellicht wat merkwaardig gekozen, maar gezien de wildgroei van speciale dagen, weken en maanden was er waarschijnlijk weinig keus.

Als thema is de waarheid natuurlijk goed gezelschap van de onderzoeker, want zijn wij daar niet allen naar op zoek? In het 701ste nummer van Hollands Maandblad twijfelt hoofdredacteur Bastiaan Bommelé of dit onderwerp bij wijsgeren wel in goede handen is: het kost hen volgens hem 'nogal veel hoofdbreken om er iets zinnigs over te zeggen'. Vervolgens fulmineert hij tegen een artikel in NRC Handelsblad van Heleen Pott, bijzonder hoogleraar Kunst en samenleving in humanistisch perspectief aan de Erasmus Universiteit Rotterdam, die de waarheid als 'relatief' voorstelt en zich beter kan vinden in de postmoderne visie dat waarheid een veranderlijke constructie is dan in de goede oude 'absolute waarheid'. Met argumenten als 'buiten de studeerkamer heeft de waarheid aan status ingeboet' en 'er bestaat een steeds grotere tolerantie voor gelul, gedraai en leugentjes om bestwil' bewijst ze haar stelling. Haar betoog kan zij, zoals Bommelé fijnjes meldt, slechts staande houden door te suggereren dat zij en de postmodernisten voor het eerst het relatieve van de waarheid inzagen – maar dan heeft zij toch enige elementaire filosofen gemist.

Historici en sociale wetenschappers zullen bij bovenstaande regels een geuw nauwelijks hebben kunnen onderdrukken. Zij wisten dit allang. Elke onderzoeker die met empirische gegevens werkt (ben je zonder data eigenlijk wel een onderzoeker?) kent de relativiteit van de waarheid.

Zelf ben ik sinds mijn studietijd verslaafd aan data. Dat begon al tijdens mijn doctoraalonderzoek in Mexico, waar wij met een groep niets anders deden dan het speuren naar datos teneinde de toen toonaangevende afhankelijkheidstheorie te bewijzen c.q. te falsificeren. Overal zagen wij dozen staan met computeruitdraaiers die toen nog het meeste op behang leken. Maar alom waren ook mannetjes die eerst stempels moesten zetten en 'autorización' riepen. Een fraaie anekdote is dat we tijdens ons onderzoek bij het Ministerio de Programacion y Presupuesto uit het Sistema Geomunicipal werden aangezien voor computerspecialisten die de plaatselijke ambtenaren kwamen instrueren, terwijl wij ons vertwijfeld afvroegen wat toch 'su consulta es por Bach' betekende. Had dat iets met de oude componist te maken? Later, veel later pas realiseerden we ons dat het om batch verwerking ging. Hoe romantisch toch allemaal maar ook vermoeiend.

DANS zal gezien de naam de romantiek van het onderzoek niet aan tasten maar het wel gemakkelijkken.

Dr. Ton van Rietbergen is docent/onderzoeker aan de Faculteit Geowetenschappen van de Universiteit Utrecht.

Gelezen

National Research Council: Expanding Access to Research Data: reconciling risks and opportunities. (Washington, 2005)

In deze studie van het Panel on Data Access for Research Purposes van de Amerikaanse National Research Council staat de verbrede toegang tot onderzoeksdata centraal met de voordelen die deze met zich brengt. Daarbij wordt ook stilgestaan bij de risico's. Het boek presenteert in vier inhoudelijke hoofdstukken een aantal aanbevelingen om aan de ene kant de toegang tot microdata voor wetenschappers en beleidsmakers te verbeteren en aan de andere kant de privacy van respondenten te waarborgen. De aanbevelingen bevinden zich op verschillende niveaus: van de basisvoorwaarden voor toegang en documentatie tot aan meer specifiek op het gebied van regionale datacentra, remote access en opleiding en training van onderzoekers. Daarnaast maakt het Panel zich sterk voor de ontwikkeling van ethische codes voor privacybescherming. Het boek is gratis online beschikbaar via <http://darwin.nap.edu/books/0309100127/html/>, maar kan ook als paperback of pdf besteld worden.

S.G.A. Smeenk, R.N. Eisinga, J.A.C.M. Doorewaard J.C. Teelken : Organisational commitment among European university employees; Documentation of a survey on predictors and consequences of organisational commitment in European universities conducted in 2004; (Den Haag, DANS, 2006; ISBN 90 6984 484 2, 72 pp.; Dans Data Guides 1)

This research report provides the documentation of a cross-sectional Web-based survey of the predictors and consequences of organizational commitment among 2325 university employees from six European countries (Belgium, Finland, Germany, the Netherlands, Sweden, and the United Kingdom). The first part of the report is a description of the research background, the sample, design and composition of the questionnaire, response rates, and representativeness of the study. The second part presents the variables, listed thematically according to the concept which they belong to. It provides for each variable the frequency distribution and some statistical measures. The third part contains the questionnaire as presented to the respondents.

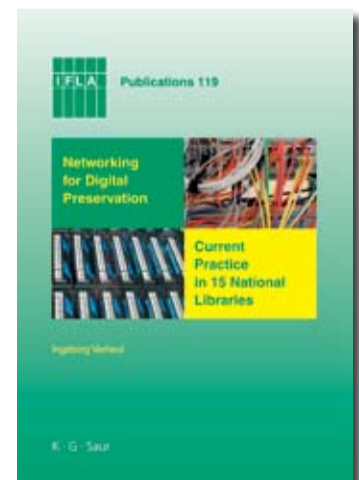
Daniel Paul O'Donnell: Cædmon's Hymn; A Multimedia Study, Archive and Edition (Cambridge, DS Brewer, 2005; ISBN 1 84384 044 8)

Cædmon's Hymn is het enig bekende werk van de zevende-eeuwse Engelse dichter Cædmon, een van de oudste poëzieteksten in de Engelse taal. De editie bestaat uit een boek en een CD-Rom. Het boek bevat een uitputtend commentaar, een weergave van alle handschriften en een reconstructie van verschillende hypothetische versies van het gedicht. De CD-Rom bevat het volledige boek en daarnaast facsimiles van alle handschriften. De tekstweergave kan worden aangestuurd door de gebruiker. Om de verbanden in de overleveringsgeschiedenis duidelijk te maken wordt inventief gebruik gemaakt van kleur en hyperlinks. Voor technisch geïnteresseerden zijn alle onderliggende coderingen op de CD-ROM beschikbaar.

Specialisten moeten beoordelen in hoeverre O'Donnell's commentaar en tekstbewerking verantwoord zijn. In technisch opzicht is de editie echter een leerzaam voorbeeld. Zelden zal een negenregelig gedicht een zo zorgvuldige behandeling zijn ten deel gevallen!

Ingeborg Verheul: Networking for digital libraries; Current practice in 15 National Libraries (München, K.G. Saur Verlag, 2006; IFLA Publications 119)

In 2004-2005 heeft de Koninklijke Bibliotheek door middel van een enquête onderzocht hoe in internationaal verband bibliotheken standaarden ontwikkelen en gebruiken op het gebied van digitale archivering. De uitkomsten van dit onderzoek zijn gepubliceerd in *Networking for digital preservation*. Lange termijn toegang tot digitale publicaties blijkt in toenemende mate een vast onderdeel te worden van het takenpakket van nationale bibliotheken. Het boek bestaat uit twee delen. In het eerste deel bevat een analyse van de activiteiten op het gebied van digitale archivering. Het tweede deel beschrijft in detail per land welke activiteiten ontplooid worden op het gebied van digitale archivering. In totaal worden vijftien landen onder de loep genomen. In toenemende mate wordt er samengewerkt om de lange-termijn toe-



gang tot digitale publicaties te waarborgen. Er is op dit moment slechts één wereldwijd geaccepteerde standaard voor digitale archivering: Het OAIS referentiemodel (Open Archival Information System). Deze standaard biedt het noodzakelijk gemeenschappelijk vocabulair om in internationaal verband met elkaar te kunnen communiceren en samen te werken.