

door Liesbeth Koenen

DANS voor data en hubs

Op zoek naar de dynamiek in de wereldongelijkheid

Jan Luiten van Zanden is verzamelaar en interpreteerder van koopkrachtplaatjes van alle tijden uit alle windstreken. Onderzoekers als hij kunnen sinds kort kennis en hulp inroepen van DANS, het goddeels virtuele instituut voor gegevensbestanden dat onder leiding staat van Peter Doorn. Reden met beiden te spreken. Een gedeeld verlangen: op wereldschaal de economische geschiedenis bij elkaar te brengen in grote bestanden, waar iedereen in de wereld mee kan werken.

Lonen en prijzen. Voor Jan Luiten van Zanden (50) zijn het de kralen waarmee hij zo'n beetje de complete wereldhistorie aaneenrijgt. Schijnbaar moeiteloos schetst de hoogleraar economische geschiedenis aan de hand van juist die basisgegevens hoe bijvoorbeeld Nederland al heel lang een kenniseconomie is, en vrouwen hun eerste serieuze emancipatiegolf beleefden vanaf halverwege de veertiende eeuw, nadat de 'zwarte dood' een kwart van

de Europese bevolking had uitgeroeid. 'De prijzen en de lonen zijn de DNA-afdruk van de economie,' stelt Van Zanden in zijn werkkamer op het Internationaal Instituut voor Sociale Geschiedenis van de KNAW in Amsterdam.

De grote vraag die hij wil beantwoorden: waarom zijn sommige delen van de wereld rijk en andere arm? 'Hoe valt te verklaren dat de afgelopen twee-, driehonderd jaar West-Europa aldoor relatief rijk is geweest, en dat pas de laatste tijd bijvoorbeeld China en India zich ontwikkelen? Je moet het op de lange termijn zien. Rond 1000 was het Midden-Oosten veel verder ontwikkeld dan Europa. Welke elementen zijn daarna dan een rol gaan spelen? Als je dat allemaal wilt weten, moet je naar de relatieve lonen en prijzen kijken. Hoeveel brood of rijst kun je kopen voor je loon? Zijn er verschillen in betaling tussen het ene beroep of het andere? Wordt er voor kennis betaald? Is die schaars of goedkoop? Verdient een scheepstimmerman bijvoorbeeld veel of juist niet? Dat zegt veel over de ontwikkelingsmogelijkheden.'

Van Zanden schiet met aanstekelijk enthousiasme door de eeuwen en de werelddelen: 'De laatste jaren gaan de archieven in China echt open. Daar zitten wij bovenop. Door de Spinozapremie die ik in 2003 kreeg, kunnen we mensen betalen die daar onderzoek doen. Het ideaal is om op wereldschaal de economische geschiedenis bij elkaar te brengen in grote bestanden, waar iedereen in de wereld mee kan werken, of het nou in China, Azië, Afrika of hier is.'



Jan Luiten van Zanden (foto Henk Thomas)

Met dat ideaal in gedachten, werkte Van Zanden vorig jaar mee aan een aanvraag voor de honderd miljoen euro die het kabinet beschikbaar stelde voor een nationaal programma ter versterking van de wetenschappelijke infrastructuur. De aanvrager was DANS, het goeddeels virtuele instituut van de KNAW en NWO samen, dat afgelopen juni van start ging. DANS staat voor 'Data Archiving and Networked Services'. Directeur dr. Peter Doorn vat het doel simpel samen: 'Onderzoekers uit de sociale en geesteswetenschappen toegang verschaffen tot digitale databestanden.' Dat hij het in één zin kan zeggen, laat

Prijzen en lonen zijn de DNA-afdruk van de economie

volgens Doorn (48) zien dat DANS een duidelijke focus heeft – in tegenstelling tot het nogal heterogene NIWI, het Nederlands Instituut voor Wetenschappelijke Informatiediensten van de KNAW, waar hij hiervoor werkte en dat inmiddels grotendeels is opgesplitst over een aantal instituten en diensten. DANS is daar een belangrijke van.

Die NIWI-achtergrond van DANS zorgt voor nogal wat scepticisme in het veld, gebiedt de eerlijkheid Van Zanden te zeggen. 'DANS zal meer naar buiten moeten kijken, een andere cultuur moeten hebben dan het NIWI,' vermaant hij. 'Maar als ze inderdaad dienstbaar aan wetenschappers kunnen zijn, en op zoek gaan naar researchgroepen om mee samen te werken, dan verwacht ik zeker voordelen van de extra expertise. We zijn hier vanuit het IISG natuurlijk al lang bezig met databestanden. Met ze verrijken, verdiepen. Maar ik ben toch maar een amateur op dat gebied.' Hij grinnikt: 'Het is geen geheim dat goed bewaren niet mijn sterkste punt is. Ik hou me niet met de organisatie van bestanden bezig...'

Rijp en rot

Dat doet DANS wel. Doorn: 'Alles is nu zo anders dan vroeger. In de jaren zeventig had je nog grote mainframecomputers. Dan bestelde je een tape. Nu kan iedereen z'n eigen data archiveren en aanbieden. Dat is heel mooi en goed, maar er is nu wel een stortvloed van groen, rijp en rot door elkaar. Nou denk ik niet dat we met een clubje als DANS – er werken nu twaalf van de geplande twintig mensen – het hele sadistisch universum van het internet

kunnen gaan ordenen, maar we kunnen wel wat bijdragen.'

Bijvoorbeeld criteria voor de betrouwbaarheid van digitale databewaarplaatsen. Maar DANS moet zelf ook een belangrijke en nuttige verzamelaarsplaats zijn, zegt Doorn. 'Een soort veiligheidsdepot, voor databestanden vergelijkbaar met het e-depot voor publicaties van de Koninklijke Bibliotheek. Die heeft niet alleen van gedrukte maar ook van elektronische publicaties altijd nog een exemplaar. Eén functie van DANS is kopietjes verzamelen van alle wetenschappelijke dataverzamelingen die er al zijn. Om de activiteiten uit te breiden werkt DANS samen met onderzoeksgroepen aan thematische ontwikkelprogramma's. Dan gaat het om onderzoekers uit een bepaald gebied die rond een grote hoeveelheid gegevens onderzoek willen doen. Er is nu bijvoorbeeld een aanvraag van taal- en tekstwetenschappers om verschillende corpora aan elkaar te knopen. Die kun je dan vervolgens zowel uit een letterkundig als een taalkundig perspectief bevragen.'



Peter Doorn (foto Wiebe Kiestra)

DANS moet voor die programma's telkens financiële partners vinden. Doorn: 'We hameren echt op het aambeeld werkt samen!' Dat geldt ook voor de tweede lijn waarlangs DANS wat structuur wil aanbrengen. 'Archivering ging vroeger een-op-een,' legt Doorn uit. 'Een medewerker van een data-archief las iets in een tijdschrift en begreep daaruit dat er data verzameld waren. Die nam dan contact op en vroeg of die gegevens in pakweg het Steinmetz-archief mochten worden opgenomen. Daar ging vaak maanden overheen, met heen en weer bellen, mailen en overleggen over het toegankelijk maken van de gegevens, et cetera. Zo kun je niet meer dan enkele tientallen datasets per jaar verzamelen. Terwijl er honderden, duizenden zijn, en er steeds meer komen. Dat moet dus anders.'

Omrekenprogrammaatje

'We proberen nu als volgt samen te werken met een instituut of een faculteit: wij stellen een medewerker ter beschikking en helpen en begeleiden, op voorwaarde dat de instelling zelf ook menskracht levert. In het kader van het NIWI hebben we zo al een proefproject met het Meertens Instituut gedaan en dat leverde met anderhalve medewerker 800 databestanden op in twee jaar tijd. We doen nu ook zo'n project met archeologen. En daardoor zijn er afgelopen jaar meer archeologische databestanden bij gekomen dan in de geschiedenis en de sociale wetenschappen. Het werkt dus onmiddellijk.'

Dat heel veel gegevens nog helemaal niet gedigitaliseerd zijn, ligt voor de hand. Die zitten verstopt in boeken en studies en archieven. Neem Van Zandens lonen en prijzen. Die hebben bijvoorbeeld geleid tot een prachtig omrekenprogrammaatje op de website van het IISG dat toegankelijk is voor iedereen. (Van Zanden: 'Met de

coste, hoeveel geld is dat in hedendaagse termen?' Of hoeveel in vergelijking tot een eeuw eerder of later? De mogelijkheden zijn eindeloos: dat salaris van vierhonderd gulden dat je vader in 1958 verdiende, blijkt een koopkracht gehad te hebben van 1053 euro nu. Een Middeleeuws dubbeltje (in guldens) was ongeveer hetzelfde als een tientje (in euro's) nu.

Maar hoe kun je dat eigenlijk weten, hoe komt Van Zanden aan al die gegevens? 'Nou, steden bouwden bijvoorbeeld vestingwallen. Er waren gasthuizen. Die huurden mensen in voor een vertimmering. Daar zijn in de archieven rekeningen van. Zo kom je erachter wat een timmerman in Leiden in 1393 verdiende. Of een spinster. Maar je hebt ook de rekeningen van een ziekenhuis of een vondelingentehuis. Daarop staat wat ze kochten, dus hoeveel het graan kostte. Op die manier krijg je ook een



Huishoudboekjes uit de zeventiende tot de negentiende eeuw (foto Henk Thomas)

Er bestaan jaarlijkse rekeningen van de werkzaamheden aan de dom van Utrecht

nieuwe kreet heet dat kennisvalorisatie – dat je een wijder publiek kennis laat maken met je resultaten.') Het vertelt wat geld in Nederland tussen 1450 en nu waard was. Je kunt alle jaren en alle bedragen onderling vergelijken door ze simpelweg in te tikken. Van Zanden: 'Dus dat huis dat familiebezit is en in 1862 zoveel honderd gulden

beeld van wat dat timmermansloon waard was. Of neem de Utrechtse dom. Er bestaan jaarlijkse rekeningen van de werkzaamheden. Daarop kun je zien dat ze leesteen uit de Hartz haalden, en wat de spijkers kostten. En als spijkers nou duurder waren in Nederland dan in China bijvoorbeeld, dan zegt dat iets over de ijzerindustrie. En dan zit je al gauw bij brandstof en de vraag hoe duur die was.'

'Veel gegevens zijn in keurige boeken gepubliceerd. Die halen we zoveel mogelijk uit de boeken en gieten we in een digitale vorm, waarmee we dan 'hubs' maken, digitale knooppunten voor gegevens. Daarvoor moet je de dingen ook vergelijkbaar maken. Hoe waren de prijzen in Krakau in verhouding tot de rest van Europa? Op zulke vragen wil je antwoord. Vaak druk je het uit in zilver. Dat is de basis van het monetaire systeem. Soms is het heel lastig, verschillen dingen van dorp tot dorp. We zijn bijvoorbeeld ook bezig met Babylon. Dan zit je met een heel andere jaarrekening. Maar goed, je bekijkt dus hoeveel mosterd of gerst kostte in termen van grammen zilver of

goud. En je kijkt hoe efficiënt zo'n markt is. Zit er veel variatie in de prijzen? Een goede markt kent weinig fluctuaties. Die heb je juist bij misoogsten, oorlogen en andere rampen. Dat zegt allemaal iets over de ontwikkeling van een economie, in dit geval de Soemerische.'

Girl power

'En in welke mate is er sprake van vrije markten? Daar gaan wij altijd wel van uit, maar in China waren er periodes zonder lonen. Belasting ging in de vorm van arbeid: dan moest je een maand komen werken in Beijing. Soms is onze markteconomie er niet. De veronderstelling is verder dat een sterk gedirigeerde markt minder ontwikkelingspotentieel heeft. Daar houden de methoden op. Maar als je de dynamiek in de wereldongelijkheid begrijpt, kun je zelfs in de toekomst kijken. We zijn bezig met een zeer experimenteel paper, waarin we proberen te kwantificeren welk deel zich ontwikkelt, welk deel achterblijft en wanneer de wereld voor honderd procent



Foto Henk Thomas

ontwikkeld is. Ons model veronderstelt dat dat kan. Dat is dus een ander uitgangspunt dan het idee dat rijkdom en armoede per se naast elkaar nodig zijn.'

Maar is dat niet erg koffiedikkijken? Ontbreken er niet per definitie te veel gegevens? Van Zanden: 'Het is speculatief, en we zijn er nog niet uit. Maar als je van 150 landen het inkomen en de bevolking kent, en statistische en demografische gegevens hebt, dan kun je ook gaan verfijnen. Je neemt bijvoorbeeld de verstedelijkingsgraad mee en de samenstelling van de beroepsbevolking, en zo verder.'

Daarom bellen alle Italiaanse mannen zodra het vliegtuig landt als eerste hun moeder

Dingen aan elkaar knopen, en zo door de geschiedenis rollen, Van Zanden doet het met verve, en benadrukt dat ook culturele patronen en de normen en waarden achter gedrag van groot belang zijn voor het verklaren van economische ontwikkelingen. Zo werkt hij nu aan een boek dat *Girl Power* moet gaan heten, en waarin de relatieve macht en vrijheid van West-Europese vrouwen sinds

de dertiende eeuw wordt verklaard. Het heeft maken met erfoverdracht, wel of geen bruidschatten, huwelijksleeftijd. Eigenlijk zie je volgens Van Zanden in de huidige verschillen tussen Zuid- en Noordwest-Europa nog steeds van alles dat terug te voeren is op de verschillen tussen Germaans en Romaans recht. Het Germaanse individualisme versus de Romaanse traditie van de ouders die bij hun kinderen blijven en andersom. Van Zanden lachend: 'Daarom bellen alle Italiaanse mannen zodra het vliegtuig in pakweg Milaan landt als eerste hun moeder.'

Goede hoop

Een superhub, een gigazoeckmachine op de website van DANS waar al die en nog veel meer gegevens doorzocht kunnen worden, dat was een van de wensen in de heel brede aanvraag waarmee DANS mee wou dingen naar de genoemde honderd miljoen voor wetenschappelijke infrastructuur. Helaas. NWO besloot uiteindelijk om vijf andere projecten in te dienen, maar het DANS-voorstel zat wel bij de laatste tien. Doorn en Van Zanden hebben allebei goede hoop dat het er alsnog een keer van zal komen.

Data Archiving and Networked Services

DANS

Het webadres van DANS luidt: www.dans.knaw.nl/

Het geld- en jarenomrekenprogrammaatje is te vinden op: www.iisg.nl/hpw