

GESPROKEN WOORD GELDT

# **EEN ACADEMISCHE LENTE**

**door**

**Dymphna C. van den Boom**

**Rede ter gelegenheid van de 384<sup>ste</sup> dies natalis  
van de Universiteit van Amsterdam  
op vrijdag 8 januari 2016**

## Een academische lente

Dames en heren,

Tim Gowers is niet wat je noemt een doorsnee blogger. Hij is wiskundige, werkt in Cambridge en ontving de Fields medal. Op zijn blog staan wiskundige ideeën en inzichten. In januari 2009 besloot Gowers zijn blog in te zetten voor een ongebruikelijk experiment. Hij selecteerde een belangrijk en lastig wiskundig probleem, één dat hij dolgraag zou willen oplossen. Maar in plaats van zelf, of in samenwerking met een paar naaste collega's, het probleem aan te pakken, besloot hij het volledig open te doen door ideeën en deeloplossingen op zijn blog te zetten. Ook plaatste hij een uitnodiging, waarin hij om hulp vroeg. Iedereen kon het proces bijhouden, een idee toevoegen of toelichten in de commentaar sectie van de blog. Gowers hoopte dat veel knappe koppen effectiever zouden zijn dan één. Dat mensen met uiteenlopende expertises en perspectieven elkaar zouden stimuleren en dat het werk aan dit lastige probleem daardoor collectief makkelijker zou zijn. Hij noemde het experiment het Polymath-project.<sup>1</sup>

Het kwam maar langzaam op gang. Zeven uur nadat Gowers zijn blog plaatste, had nog niemand gereageerd. Toen kwam er een commentaar van Solymosi, wiskundige aan de Universiteit van British Columbia. Een kwartier later volgde Jason Dyer, een leraar uit Arizona, met een ander idee. Weer drie minuten later plaatste Terence Tao, een wiskundige uit UCLA, en eveneens winnaar van de Fields Medal, zijn visie. En vervolgens barstten de commentaren los: in 37 dagen plaatsten 27 personen 800 berichten. Bij lezing van al die commentaren valt op dat er met een ongelooflijke snelheid ideeën worden geopperd, verfijnd en weer verworpen. Je ziet topwiskundigen fouten maken, verkeerde paden inslaan en hun handen vuil maken aan de simpelste details. En tussen alle valse starten en verkeerde afslagen door, begint geleidelijk aan het inzicht te gloren.<sup>2</sup>

Zevenendertig dagen na de start van het project was niet alleen het probleem opgelost, maar ook een nog lastiger theorema, waarvan het oorspronkelijke een speciaal geval was. Gowers beschreef het hele experiment als 'één van de opwindendste zes weken in zijn loopbaan'. Sinds dit eerste project zijn er al een tiental polymath-achtige projecten gelanceerd.<sup>3</sup>

Wat Gowers deed is het efficiënt verbinden van de denkkraft van een grote groep deskundigen. Het is een prachtig voorbeeld van wat Open Science vermag. Open Science is een breed concept: het omvat open educational resources, open access, open peer review, open methodologie, open source en open data. De term werd in 2003 gelanceerd door de

---

<sup>1</sup> M. Nielsen. *Reinventing discovery: The new era of networked science*. Princeton NJ: Princeton University Press, 2011 (p.1)

<sup>2</sup> Idem (p.1)

<sup>3</sup> Idem (p.2)

econoom Paul David<sup>4</sup> om een beweging te beschrijven die wetenschappelijk onderzoek, onderzoeksdata en de disseminatie ervan toegankelijk maakt voor een leergierige gemeenschap van amateurs tot professionals.

### Een tweede revolutie

De opkomst van wetenschappelijke tijdschriften wordt wel de eerste Open Science-revolutie genoemd. Vóór de 17<sup>e</sup> eeuw werden wetenschappers door de maatschappij tot een zekere mate van geheimzinnigheid gedwongen. Zij hadden weinig te winnen, maar veel te verliezen met het bekendmaken van hun wetenschappelijke doorbraken. Dat veranderde met de opkomst van de academies en de tijdschriften. Engeland richtte in 1660 de Royal Society op, en 6 jaar later de Fransen de Academie van Wetenschappen. Zo'n 70 andere wetenschappelijke organisaties kregen daarna ook officiële erkenning. In 1665 werd Henry Oldenburg editor van *Philosophical Transactions* van de Royal Society, het eerste tijdschrift volledig gewijd aan de wetenschap. Het vormde de basis voor de explosieve groei van de wetenschappelijke uitgeverij. Eind 17<sup>de</sup> eeuw waren er 30 wetenschappelijke tijdschriften; rond 1790 waren er al 1052. En sindsdien is het publiceren alsmaar toegenomen.

We staan nu aan de vooravond van een tweede Open Science-revolutie<sup>5</sup>, die zeker zo groot zal zijn als de eerste. De veranderingen die deze met zich mee zal brengen, raken het hart van de wetenschap en reiken veel verder dan de eis om te publiceren of data openbaar te maken.

Open Science verandert op de eerste plaats de wijze van samenwerken. Ten tweede verandert het fundamenteel de traditionele methode van wetenschap bedrijven. Ook beïnvloedt het de relatie tussen wetenschap en maatschappij. Ik zal successievelijk op deze drie aspecten ingaan.

De Polymath Projecten illustreren het eerste aspect, namelijk hoe wetenschappers op een andere manier kunnen samenwerken. Ze laten zien dat nieuwe communicatietechnologieën de potentie hebben een nieuwe sociale dynamiek in de wetenschap te creëren, áls er de bereidheid is elke stap in het proces te delen en de traditionele werkwijze wordt losgelaten.

De tweede verandering is gerelateerd aan de spectaculaire groei van de hoeveelheid beschikbare data. Computers maken het mogelijk enorme datasets te creëren en te exploreren op manieren die inherente, maar onvermoede relaties blootleggen. Een voorbeeld hiervan is Don Swanson, een gepensioneerde, maar nog actieve informaticus. Hij gebruikte Medline, een database van miljoenen medische artikelen en deed verschillende ontdekkingen zonder ooit een medische opleiding genoten te hebben. Zo vond hij een verband tussen migraines en magnesium door naar relaties te zoeken tussen migraines en andere condities. Geen enkel

---

<sup>4</sup> Paul A. David. Understanding the emergence of 'open science' institutions: functionalist economics in historical context. *Industrial and Corporate Change*, 13, 4, pp. 571-589.

<sup>5</sup> *Science as an open enterprise*. The Royal Society Science Policy Centre Report 02/12.

individue kan ooit de miljoenen onderzoeken in Medline bevatten, maar één individu kan er wel patronen in ontdekken, waarvan het bestaan niet kon worden vermoed.

Een ander voorbeeld is het gezamenlijk in kaart brengen en verzamelen van data; zoals in de GenBank<sup>6</sup> door biologen, of in HapMap<sup>7</sup>, voor de genetische variaties tussen individuen. Op zeer veel andere terreinen werken wetenschappers online samen om zulke databestanden te creëren. Bestanden die de structuur van het heelal, het klimaat wereldwijd of alle flora en fauna in kaart brengen. Met behulp van deze geïntegreerde databestanden kan iedereen algoritmes gebruiken om naar verbanden te zoeken, waarvan men het bestaan niet eens bevroedde.

Sommigen noemen dit een vierde paradigma in de wetenschap<sup>8/9</sup>. Aan het klassieke duo van experiment en theorie werd na de komst van de moderne computer een derde toegevoegd: de simulatie. Het vierde paradigma zet alles op zijn kop. In plaats van het ontwikkelen en toetsen van hypothesen uit data die voor dat doel verzameld zijn, worden hypothesen geconstrueerd, *nadat* relaties in de dataset geïdentificeerd zijn. En dat op een immens veel grotere schaal dan tot op heden mogelijk was. Het zet niet alleen de methode van onderzoek op zijn kop, maar zorgt ook voor een aardverschuiving in mogelijke bevindingen. Vandaar een vierde paradigma.

De derde grote verandering als gevolg van Open Science betreft de relatie tussen wetenschap en maatschappij. Een voorbeeld is de website Galaxy Zoo, die meer dan 200.000 online vrijwilligers gerekruteerd heeft om astronomen te helpen melkwegstelsels te classificeren. Galaxy Zoo kan gezien worden als een kosmologische telling, de grootste die ooit ondernomen is. Een telling die tot op heden meer dan 120 miljoen classificaties heeft opgeleverd en 25 wetenschappelijke publicaties gebaseerd op data van Galaxy Zoo. Andere voorbeelden zijn Foldit, een computerspel dat vrijwilligers online uitdaagt, uit te zoeken hoe aminozuren proteïne creëren en de applicatie *Vogel het uit*, ontwikkeld door onderzoekers van het Instituut voor Biodiversiteit en Ecosysteem Dynamica (IBED) van de UvA, waarbij met behulp van vogelspotters vragen beantwoord werden over het leven en het gedrag van vogels. IBED won met de app de Academische jaarprijs in 2013. De potentie van deze zogenaamde *citizen science* is natuurlijk nog ongewis. Maar zeker is dat de groei van de beweging tot een andere dynamiek in de wetenschap kan leiden, waarbij de scheidslijn tussen professional en amateur vervaagt en in elk geval de aard van de betrokkenheid van het publiek bij de wetenschap verandert.

---

<sup>6</sup> D.A. Benson, M. Cavanaugh, K. Clark, I Karsch-Mizrachi, D. J. Lipman, J. Ostell & E. W. Sayers (2013). GenBank. *Nucleic Acid Research*, 41, Database issue.

<sup>7</sup> The International HapMap Consortium (2003). The International HapMap Project. *Nature*, 426, 18/25 december.

<sup>8</sup> J. Gray (2009). *E-Science: a transformed scientific method*. In: The fourth paradigm: data-intensive scientific discovery. Hey, T., Tansley, S., & Tolle, K. (Eds.). Microsoft Research: Washington.

<sup>9</sup> Shadbolt, N, Berners-Lee, T., & Hall, W. (2006). *The semantic web revisited*. IEEE Intelligent Systems, 21, 3, 96-101. Available at: <http://eprints.soton.ac.uk/262614/>.

Ik besluit met een voorbeeld waarin het allemaal samenkomt: Open Science in real time tijdens een uitbraak van een ernstige darminfectie in Hamburg in 2011<sup>10</sup>. Deze verspreidde zich in verschillende landen in Europa en de VS, trof zo'n 4000 mensen en leidde tot 50 sterfgevallen. Alle patiënten testten positief voor een onbekende variant van de *E. coli* bacterie. De stam werd aanvankelijk geanalyseerd in Shenzhen in China in samenwerking met Hamburg en 3 dagen later werd er een nog niet definitief bevestigd genoom bekend gemaakt onder een open data-licentie. Dit trok de aandacht van bioinformatici uit vier werelddelen. Vierentwintig uur na het vrijgeven was het genoom geassembleerd. Binnen een week stonden er 24 rapporten op een open source-site, gewijd aan de analyse van de bacterie stam. De resultaten verschenen op tijd om de uitbraak onder controle te houden. Het is een indrukwekkend voorbeeld van een nieuwe manier van samenwerken, in combinatie met open data en het delen van databestanden.

### **Progressieve doelen, conservatieve middelen**

Tot nu toe presenteerde ik u een optimistisch verhaal over Open Science, maar er zijn ook obstakels, die ik nog verzwegen heb. Het delen van data is, op dit moment in elk geval, een merkwaardige keuze, want als wetenschapper is het in je voordeel je data zo lang mogelijk geheim te houden. Waarom zou je online je data delen, voordat je zelf de kans gehad hebt er een artikel van te maken? Waarom zou je je hypothese delen, zodat iemand anders de analyse kan maken? Waarom zou je een denkrichting lanceren, met het risico iemand anders op weg te helpen?

In de wetenschappelijke wereld zijn het publicaties, die leiden tot banen en promotie. Zo is de academie georganiseerd. Het publiek ter beschikking stellen van data draagt bepaald niet bij aan je loopbaan en kan die zelfs belemmeren door het helpen van de concurrent. Een belangrijke barrière voor het algemeen aanvaarden van de principes van Open Science liggen dus in het systeem van beloning, waardering en promotie binnen universiteiten en onderzoeksinstituten.

Het is daarom cruciaal dat het genereren van belangrijke datasets, het beheer ervan en de open wijze van wetenschap bedrijven erkend en beloond worden. De vaardigheid en creativiteit die nodig is om op een succesvolle manier grote datasets te creëren, moeten ook als zodanig erkend worden. Citaties van open data moeten gelijk gesteld worden aan die van een conventionele wetenschappelijke publicatie. Ook nieuwe mogelijkheden als science wikis, science blogging en open peer review maken een herbezinning noodzakelijk. Altmetrics<sup>11</sup> kunnen daarbij helpen – het is een variëteit aan manieren om de impact van onderzoek te meten die verder gaat dan citaties in tijdschriften, en onder andere ook het open gebruik en delen van data meetellen. Alhoewel wetenschap progressieve doelen heeft,

---

<sup>10</sup> World Health Organization (2011). *Outbreaks of E. coli 0104:H4 infection*. Available at: <http://www.euro.who.int/en/health-topics/disease-prevention/food-safety/outbreaks-of-e.-coli-0104h4-infection>.

<sup>11</sup> R. Kwok (2013). Altmetrics make their mark. *Nature*, 500, 22 augustus,.

gebruikt het vaak nog conservatieve middelen, terwijl verbeteringen in het proces van wetenschap bedrijven belangrijker zou moeten zijn dan individuele ontdekkingen.

## Een transformatie in de wetenschap

Dames en heren,

Genoemde voorbeelden zijn slechts een klein onderdeel van een veel groter verhaal, een verhaal over de wijze waarop Open Science het wetenschappelijke proces transformeert. Zoals de eerste Open Science-revolutie het wetenschappelijke proces geleidelijk transformeerde, zal de tweede revolutie het bedrijven van wetenschap in de komende twintig jaar radicaal veranderen.

Er zal een wetenschappelijke cultuur moeten komen, waar zoveel informatie als mogelijk open toegankelijk wordt. Het gaat dan om *alle* informatie die van wetenschappelijke waarde is: van ruwe data tot computer codes, tot alle vragen, ideeën en speculaties die nu niet toegankelijk zijn.

Dat betekent een grote cultuuromslag. Het belang van de eerste Open Science-revolutie was om het individuele en het collectieve belang op elkaar af te stemmen door wetenschappers te belonen voor het delen van hun resultaten in wetenschappelijke tijdschriften. Hoewel het nu in het collectieve belang van wetenschappers is om nieuwe technologieën te omarmen, is het nog steeds in het individuele belang zich te richten op publicaties. Dat individuele belang zou weer moeten worden omgezet in een collectief belang. Wetenschappers kunnen dat niet alleen. Verschillende instanties zullen daarin een rol moeten vervullen. Iedereen moet zich aanpassen.

Academies moeten wereldwijde maatstaven voor open data en metadata creëren. Universiteiten zullen hun HRM-beleid moeten aanpassen, zodat de vernieuwingen die Open Science genereert, gewaardeerd worden. Deze aanpassingen zijn ook noodzakelijk bij onderzoeksvisitaties. Wetenschappers zullen bereid moeten zijn hun data open ter beschikking te stellen, uiteraard na een periode van exclusieve toegang tot de eigen data en met inachtneming van privacy en veiligheid<sup>12</sup>. De overheid zal beleid moeten maken voor het ter beschikking stellen van onderzoeksdata. Subsidiegevers moeten dat ook doen. Zij kunnen ook de handen ineenslaan bij het steunen van de ontwikkeling van de benodigde tools, software en personeel dat vitaal is voor het welslagen hiervan. Dit lijkt mij een thema uit de Nationale Wetenschapsagenda dat met de hoogste prioriteit zou moeten worden opgepakt, zowel inhoudelijk als financieel. Alle universiteiten in Nederland hebben daar belang bij, en met het juiste beleid en middelen kan Nederland op dit gebied een voortrekkersrol spelen.

---

<sup>12</sup> OECD (2015). Making open science a reality. *OECD Science, Technology and Industry Policy Papers*, No. 25, OECD Publishing, Paris. <http://dx.doi.org/10.1787/5jrs2f963zs1-en>.

We hebben nu de kans om de wijze waarop kennis gecreëerd wordt te veranderen. Investeren in deze exemplarische route in de Nederlandse Wetenschapsagenda kan een tweede Open Science revolutie in Nederland ontketenen.

Dank voor uw aandacht.