

Bouwen aan een wiskundig web

Enigzins verrast was ik wel, toen Peter de Paepe mij vroeg een stukje te schrijven over mijn belevenissen na mijn afstuderen in de (zuivere) wiskunde, bijvak Russisch. En als het kan “iets over elektronisch uitgeven erin” schreef hij erbij, wetende dat ik voor een wetenschappelijke uitgeverij werk.

Mijn gedachten gingen terug naar de bibliotheek van de subfaculteit Wiskunde, en in het bijzonder naar de IBM-schrijfmachine (met verwisselbare kop, zodat een aantal wiskundesymbolen en het Griekse alfabet beschikbaar waren). Ik heb er aan het eind van en kort na mijn studie voor diverse uitgeverijen nog enkele korte vertalingen op getypt (Russisch→Engels, voornamelijk tijdschrift-artikelen). Een langere vertaling (een boek) schreef ik indertijd gewoon uit en leverde ik als handgeschreven manuscript in. Later gebruikte ik een Apple met daarop T_EXtures. Aanvankelijk stuurde ik uitdraaien op, soms vergezeld van een bestand op floppy. Met de opkomst van email werd de procedure meer gestroomlijnd: het was voldoende het bronbestand per email te sturen, mits gecodeerd volgens aanwijzingen van de betreffende uitgever. Proeflezen was dan ook niet meer nodig, een enorme tijdwinst! Na de val van de Muur wordt veel van het weinige overgebleven vertaalwerk in de (voormalige) Soviet Unie zelf gedaan; indien noodzakelijk wordt een manuscript opgelapt door een Engelstalige expert.

Gelukkig viel niet al het werk weg. Het toeval wilde dat vlak voor mijn afstuderen (1985) uitgeverij Reidel de uitgaverechten had verworven voor een Engelse, geannoteerde vertaling van de vijfdelige *Soviet Encyclopaedia of Mathematics*. Na een zeer kort gesprek en een lange proefvertaling werd ik aangenomen om de Managing Editor, Prof. M. Hazewinkel, hierin te assisteren. Het idee was om een groep wiskundig geschoolde vertalers de vertalingen te laten maken, en deze vertalingen wetenschappelijk-wiskundig te laten controleren en becommentariëren door experts. Er was

een toen ongebruikelijke methode van publiceren gekozen: de artikelen werden gecodeerd opgeslagen op een UNIX mainframe en een aangepaste versie van troff/eqn werd gebruikt voor publikatie (dit alles op het CWI, waar men hiermee al enige ervaring had). Mijn taak was een en ander te coördineren en zorg te dragen voor consistentie en publikatie. In concreto: veel administratie en veel proeflezen.

Nadat de tien delen met geannoteerde vertalingen verschenen waren, is besloten dit project “levend” te houden, door wiskundigen uit te nodigen nieuwe artikelen bij te dragen. Het mainframe werd vaarwel gezegd, we stapten over op pc, LINUX en LaTeX. Het runnen van dit project kreeg nieuwe kanten, inhoudelijk onder meer: het vinden van titels voor nieuwe artikelen (dat wil zeggen het enigszins bijhouden van stromingen in wiskundeonderzoek), het uitnodigen van wiskundigen een artikel bij te dragen, (her)codering van ingezonden artikelen (inclusief het koppelen aan reeds bestaande artikelen), (her)schrijven van conversie-, layout- en indexeringssoftware. Tegelijk met publikatie van het eerste Supplement verscheen een CD-ROM met alle artikelen en talloze, simpele, animaties en 3D-objecten. Inmiddels zijn drie Supplementen (ieder met eigen index) verschenen.

En nu? Doorgaan met publiceren van Supplementen is een mogelijkheid, maar leidt tot diverse problemen, zoals: om een zoekterm te vinden, moet een gebruiker nu vier indexen raadplegen, een nieuw artikel kan alleen verwijzen naar reeds bestaande artikelen – in zekere zin verzwakt dit de bestaande verwijsstructuur. Een andere mogelijkheid is herdruk van de set, met nieuwe artikelen geïntegreerd. Dit stuit op andere problemen, onder meer: wie zou zijn bestaande set willen vervangen vanwege voornamelijk betere zoek- en verwijsmogelijkheden, en tegen welke prijs? Het lijkt ons beter de weg te volgen die met de CD-ROM is ingezet, die van elektronisch uitgeven.

Nu heeft de term *elektronisch uitgeven* verschillende invullingen: het bestellen en produceren van boeken via printing-on-demand (dat wil zeggen afdrukken en bezorgen per bestelling in plaats van levering uit voorraad), het vereenvoudigen van aanlever- en review-

processen voor tijdschriftartikelen (via email of een preprintserver, inclusief online status tracking), het aanbieden en toegankelijk maken van materiaal (bijvoorbeeld een tijdschrift) via een CD-ROM, ftp- of website. Bij het aanbieden van materiaal wordt in eerste instantie gedacht aan verspreiding: een artikel wordt bijvoorbeeld als afdrukbaar PDF-bestand aangeboden, of een boek wordt vergezeld van een CD-ROM met afbeeldingen. Echte hypertextproducten zijn nog zeldzaam.

Uitgeverijen zijn bezig het elektronisch uitgeven in hun processen te integreren. Sleutelwoorden daarbij zijn vind- en (her)bruikbaarheid van de informatie.

In navolging van het URL-systeem voor websites wordt in het CrossRef-project gewerkt aan een uniek identificatiesysteem voor alle publikaties (denk aan een verfijnd ISBN/ISSN). Belangrijk hierbij is onder andere dat een verwijzing altijd tot de betreffende publikatie leidt (zodat het probleem van verdwenen of veranderde URLs ondervangen wordt). Dit is één kant van vindbaarheid. Een andere kant is het zoeken *in* informatie, hetgeen samenhangt met het hergebruik en de opslagcodering er van.

Codering in \TeX (het formaat waarin welhaast iedere Westerse tegenwoordig zijn/haar artikel of boek produceert) levert vaak een visueel redelijk ogend manuscript op. De onderliggende codering variëert echter van auteur tot auteur, en is veelal op effect gericht (“maak vet”) in plaats van op functie (“benadruk”). Wetenschappelijke uitgeverijen hebben daarom een andere, zeer gestructureerde, opslagcodering gekozen (XML). Helaas zijn er beperkte mogelijkheden voor automatische conversie van algemene \TeX naar deze opslagcodering.

Het beschikbaarstellen van informatie in een vorm anders dan op papier zodanig dat de informatie niet alleen leesbaar is maar ook bruikbaar wordt, brengt een verschuiving van layout (waarover minder controle kan worden uitgeoefend) naar ontsluitbaarheid (traditioneel het terrein van bibliothecarissen). De grote investeringen (mens-tijd, apparatuur) hierin hebben mede als gevolg dat het verkoopmodel van uitgeverijen zal (moeten) veranderen. In plaats van een abonnement op een individueel tijdschrift kan men

denken aan een pay-per-view systeem of juist een abonnement op een (groot) pakket van gerelateerde informatie.

Een stap ná elektronisch uitgeven is naar mijn mening die van interactief uitgeven: aan een bestaand product wordt op controleerbare wijze informatie toegevoegd. Een voorbeeld. De bestaande referentiestructuur in artikelen werkt in één tijdsrichting: een 1991-correctie van een 1989-artikel kan refereren aan het artikel, maar nooit omgekeerd. Dit probleem is oplosbaar door informatie (bijvoorbeeld een verwijzing) toe te voegen aan het oorspronkelijke artikel. Uiteraard mag de tekst van het oorspronkelijke artikel niet veranderd worden, maar wellicht wordt een waarschuwingshyperlink “in de kantlijn” acceptabel? Dit kan uitgebreid worden: als artikel 1 verwijst naar artikel 2 (in de referentielijst), voeg dan aan artikel 2 een verwijzing toe naar artikel 1 (in een aparte sectie). Gezien de hoeveelheid tijdschriftartikelen en de muren tussen tijdschriftredacties, uitgeverijen en wetenschappers, is de implementatie hiervan nog een aardig probleem.

De invulling van interactief uitgeven die ons bij de Encyclopaedie voor ogen staat is als volgt. Het hoofddoel van (het gebruik van) de elektronische Encyclopaedie is het snel vinden van betrouwbare informatie van het type *wat is een ... of waar gaat ... over* of *is er iets bekend over het gebruik van ... in ...*, met verwijzingen naar gerelateerde informatie. In de papieren versie zijn dit verwijzingen naar artikelen en boeken, en CrossRef zal dit elektronisch mogelijk maken. Het is tevens niet moeilijk de structuur van kruisverwijzingen binnen de Encyclopaedie zichtbaar te maken. Er is echter meer waarnaar verwezen kan worden: (privé)collecties van wiskundige objecten (3D-objecten, animaties), specialistische wiskundig georiënteerde websites (bijvoorbeeld van biografieën), lokale initiatieven op universiteiten, etcetera. Hier zou de Encyclopaedie een aanvulling (verbetering) kunnen zijn op de “lijstjes van websites met interessante wiskundige inhoud” die momenteel de ronde doen. De mogelijkheid notities bij artikelen te maken kan er enerzijds voor zorgen dat informatie niet snel gedateerd raakt, en is anderzijds een bron voor nieuwe artikelen.

Echter, veel is afhankelijk van de medewerking van wiskundigen

zelf. Enerzijds is dit inhoudelijke medewerking: bij het schrijven van artikelen, het geven van aan- en opmerkingen, het gebruik van de Encyclopaedie in elektronische publicaties, etcetera. Anderzijds is begrip nodig voor het feit dat het structureren en functioneel presenteren van informatie tijd en geld kost, en derhalve ook geld moet opleveren. De Encyclopaedie verschijnt dit voorjaar op Internet, en komend jaar zal ik werken aan optimalisatie van de onderliggende codering en het koppelen van de Encyclopaedie aan externe informatie, met name aan databases van abstracts zoals Mathematical Reviews en Zentralblatt.

Rob Hoksbergen
Rob.Hoksbergen@wkap.nl