

‘Studenten zoeken niet in boeken’

University of Michigan laat Google 7,4 miljoen titels inscannen

Het is controversieel, maar daar trekken Harvard, Oxford en Michigan zich niets van aan. Google mag miljoenen boeken inscannen. „Dit had ons 1.400 jaar gekost.”

Door onze correspondent
FREEK STAPS

ANN ARBOR, 25 NOV. John Wilkins kamer op de achtste verdieping van de monumentale hoofdbibliotheek van de University of Michigan kent vele geheimen. Eén ervan wil hij wel prijsgeven. Op het beeldscherm van de bibliothecaris, verantwoordelijk voor het controversiële inscannen van 7,4 miljoen boeken om ze daarna gratis toegankelijk te maken via internet, staat ‘The Google workstation. CONFIDENTIAL’.

Vertrouwelijk is het. Toen Google-bestuurder Larry Page in 2002 met Wilkin kwam praten om de boeken van zijn alma mater in te scannen voor Google Print, moest de bibliothecaris voor geheimhouding tekenen. Tot de bekendmaking eerder dit jaar hield Wilkin zijn mond – zelfs zijn collega’s wisten van niets. Toen Google zijn scanapparaten aan het testen was en constant dozen het gebouw in en uitgingen, liet Wilkin bordjes maken alsof het om de gebruikelijke digitalisatie van de boeken van de universiteit ging. Iedereen trapte erin.

Het was niet voor het eerst dat de universiteit benaderd werd door een bedrijf om de uitgebreide collectie van een van Amerika’s beste universiteiten in te scannen. Een grote uitgever had eens contact opgenomen, een groot internetbedrijf ook. Wilkin wil niet zeggen om wie het gaat. Maar geen van die partijen kon aan zijn drie eisen voldoen. Eén: de digitalisatie moet ‘non-destructive’ zijn, de boeken mochten er niet onder lijden. Twee: de kwaliteit moet zo hoog zijn dat de universiteit de bestanden kon gebruiken voor het eigen archief. Drie: de bibliotheek behoudt de rechten van de bestanden.

De ene partij wilde het alleen doen als ze de rechten mocht aankopen, de ander deed proeven met resultaten van onvoldoende kwaliteit. Maar dat iemand de universiteit gelukkig moest maken, dat was Wilkin wel duidelijk. De eigen digitalisatie verliep met 5.000 boeken per jaar te langzaam: „Om deze collectie in te scannen hebben wij 1.400 jaar nodig. Google kan het in zes.”

Dat wil zeggen: dat was het plan.



De centrale opslagruimte van de bibliotheek van de University of Michigan. (Foto AP/The Detroit News, Mandy Wright)

Maar het gaat voorspoediger dan verwacht, Google denkt nu al binnen drie jaar klaar te zijn. De University of Michigan en de negentien bibliotheken van de universiteit zijn daarmee het epicentrum van Google’s project. Hier probeerde het bedrijf de speciaal gebouwde scanners uit en hier wordt veruit het grootste aantal boeken gedigitaliseerd.

Hoe het in de praktijk werkt? De bibliotheken van de universiteit hebben een vestiging buiten de universiteitsstad Ann Arbor, in de noordelijke staat Michigan. Daar bestellen de Google-werknemers grote pakketten boeken, gesorteerd op de grootte van het boek. Per e-mail worden gegevens over de auteur, de titel en de inhoud naar Google gestuurd. Binnen acht dagen moeten de boeken dan weer bij de bibliotheek terug zijn, want „als Google een langere verwerkingstijd heeft, moeten ze zelf een bibliotheek bouwen”.

Dat is nog helemaal geen gek idee. Lokale media volgen de digitalisering op de voet en melden dat makelaars in opdracht van het Californische bedrijf op zoek zijn naar werk- en woonruimte voor

duizend Google-werknemers.

De onderneming wil er niets over kwijt. Het hele project blijft daarmee net zo geheim als geruchtmakend, want een storm van protest is opgestoken. Schrijversorganisatie Authors Guild, met achtduizend leden, en vijf grote Amerikaanse uitgevers, waaronder Simon & Schuster en McGraw-Hill, hebben rechtszaken tegen Google aangespannen. Zij zeggen dat Google met het toegankelijk maken van de boeken auteursrechten veronachtzaamt en dat Google tevoren toestemming had moeten vragen.

Google maakte in reactie op het protest ingescande boeken een maand later toegankelijk, maar heeft begin deze maand gewoon de eerste duizenden – ook die exacte aantallen zijn onbekend – werken doorzoekbaar gemaakt via print.google.com. Het gaat hier wel om boeken ‘uit het publieke domein’. In de praktijk betekent dat boeken waarvan de overheid de uitgever is, of die voor 1923 zijn verschenen.

Wilkin verwachtte wel protest, maar toont zich niet onder de indruk van de argumenten van de schrijvers en uitgevers. „Kopiëren

mag, kijk maar naar eerdere zaken over cassette- en videobanden. De kwestie draait om wat je met die kopieën dóét.” Dat sommigen bang zijn inkomsten te missen, doet Google af als niet terecht. Volgens het bedrijf krijgen werken zo juist meer aandacht. Bibliothecaris Wilkin: „Hiermee heb ik niets te maken. Hiervoor moeten de schrijvers bij Google zijn.”

Google maakt het – nog – niet mogelijk hele boeken te downloaden, de grootste vrees van de tegenstanders. Een zoekresultaat geeft het woord, een paar zinnen ervoor en erna en informatie over het werk. Dat spoort met Google’s missie om ‘alle informatie van de wereld te organiseren en beschikbaar te maken’ en het ligt ook in de lijn van Wilkins bibliotheek. „Ons doel is kennis te vergaren, bewaren en verspreiden.”

En daarvoor is geen beter middel dan internet, denkt de universiteit. „Want laten we eerlijk zijn: studenten van nu zoeken niet in boeken naar informatie. Ze zoeken op Google. Laten we niet klagen dat ze het verkeerd doen, maar laten we de informatie daarheen brengen waar ze toch al zoeken.”

Oh, en de afbeelding van de scanmachines op Wilkins computer? Hij klikt wat vensters weg zodat de hele foto zichtbaar wordt. Op zijn scherm verschijnt een oude televisie en een monstrueus kopieerapparaat die met een dikke kabel verbonden zijn aan een typemachine.

Boeken worden bits en bytes

• **Google:** De universiteiten Harvard, Stanford, Michigan en Oxford en de New York Public Library laten grote delen van hun collecties door Google inscannen.

De bibliotheken verdienen er niets op, maar krijgen van Google de elektronische versies voor hun archief.

• **Amazon:** gaat pagina’s van boeken los te koop aanbieden
• **Microsoft:** gaat via MSN 100.000 gedigitaliseerde boeken aanbieden. De 100.000 werken komen uit de British Library. De bibliotheek krijgt 2,5 miljoen dollar.
• **Yahoo:** bouwt samen met uitgevers en auteurs een online-bibliotheek op openlibrary.org