

## Google en de digitalisering van boeken

Google Books  
Google Books voor bibliotheken  
Google Scholar  
Amazon  
Open Content Alliance

### Google en de digitalisering van boeken

Boeken migreren naar internet: er is sprake van een golf van initiatieven om boeken te digitaliseren en beter toegankelijk te maken. Het begon in 2004 met Google, maar sindsdien hebben ook Amazon, Yahoo met partners in de Open Content Alliance en MSN Search zich gemeld. Een kleine inventarisatie.

#### Google Print

In het Google Books programma ([books.google.com](http://books.google.com)) biedt Google aan boeken gratis te digitaliseren en op te nemen in de Google-zoekmachine. Het programma richt zich op alle boeken die uitgeverijen en een aantal wetenschappelijke bibliotheken daarvoor beschikbaar willen stellen. Het aantal te digitaliseren boeken kan oplopen tot meer dan 15 miljoen. Google, dat inmiddels meer dan 8 miljard webpagina's regelmatig indexeert, kijkt van zulke aantallen niet op.

Google Books is onderdeel van de gewone Web Search, de algemene zoekmachine van Google, maar ook apart te gebruiken via Google Book Search. Mensen die via Google op zoek zijn naar informatie kunnen voortaan in de zoekresultaten naast webpagina's ook pagina's uit boeken vinden. Een link naar een boek leidt naar een plaatje van de betreffende pagina, naar gegevens over het boek, een plaatje van het omslag en links naar bestelpagina's, onder meer op de sites van Amazon, Barnes & Noble en de website van de uitgever.

De gebruiker kan vanuit de gevonden pagina enkele pagina's vooruit en achteruit door het boek bladeren en ook in het boek zoeken op een nieuw trefwoord. Vanuit nieuwe pagina's met treffers kan vervolgens weer gebladerd worden. Met deze functies probeert Google de ervaring van het snuffelen in boeken in de boekwinkel na te bootsen. De mogelijkheden om te bladeren zijn echter beperkt en het is vanwege de auteursrechten ook niet mogelijk om pagina's te kopiëren of te printen.

Google Books is een interessant marketingprogramma via het web. Uitgeverijen kunnen de zichtbaarheid van boeken vergroten, een nieuw publiek bereiken, de verkoop van boeken verhogen, het bezoek aan de eigen website stimuleren en meedelen in de inkomsten uit de advertenties die Google bij de boekpagina's plaatst.

Het programma werkt inmiddels voor een aantal talen (Engels, Duits, Frans, Spaans en ook Nederlands).

#### Google Books voor bibliotheken

Google richt zich ook op bibliotheekcollecties. In het Google Books-programma komen de complete collecties van de bibliotheken van de universiteiten van Michigan en Stanford (respectievelijk 7 en 8 miljoen titels) en delen uit de collecties van de universiteit van Harvard (40.000 van de 14,7 miljoen publicaties), de openbare bibliotheek van New York (een selectie van de 20 miljoen rechtenvrije banden) en de universiteit van Oxford in Engeland (ongeveer 1 miljoen titels van vóór 1900).

De copyrightvrije boeken zullen integraal via internet beschikbaar worden gesteld. Van boeken waarop nog auteursrechten rusten zullen maximaal drie korte passages met treffers en de bibliografische gegevens worden getoond. Voor zover bekend delen de bibliotheken niet in de advertentie-inkomsten op 'hun' pagina's, maar ontvangen ze wel de bestanden met gedigitaliseerde boeken uit hun collectie en kunnen deze ook via hun eigen sites beschikbaar stellen.

Dit programma is op veel weerstand gestuit, omdat veel uitgevers en auteurs van mening zijn dat het digitaliseren van boeken waarop auteursrecht rust toestemming vereist van de rechthebbende. Deze kwestie, die draait om de vraag wat *fair use* precies inhoudt, komt nu voor de Amerikaanse rechter.

#### Google Scholar

Google Scholar ([scholar.google.com](http://scholar.google.com)) is een voor iedereen toegankelijke gespecialiseerde zoekmachine voor wetenschappelijke informatie, in de vorm van artikelen, rapporten, recensies, boeken en manuscripten. Treffers zijn gerangschikt naar relevantie, onder meer door het aantal literatuurverwijzingen naar het betreffende document te tellen. Via een link ('Cited by') kunnen deze documenten worden opgeroepen. Boeken kunnen vervolgens worden opgespoord door te klikken op 'Web Search', waarna de algemene Google zoekmachine naar het betreffende boek zoekt, of op 'Library Search', die zoekt in een naburige bibliotheek. Als het betreffende boek ook is opgenomen in Google Books kan het direct worden opgeroepen.

Uitgeverijen en instellingen die hun wetenschappelijke informatie willen laten indexeren voor Google Scholar kunnen zich eenvoudig opgeven. Google stelt als voorwaarde dat bezoekers minimaal toegang

krijgen tot samenvattingen van publicaties en de opgenomen literatuurverwijzingen.

Google Scholar zoekt echter niet alleen in publicaties van traditionele wetenschappelijke uitgeverijen, maar ook in openbaar toegankelijke bronnen op het web, in het bijzonder 'open access' publicaties en de 'repositories', digitale archieven van universiteiten en wetenschappelijke instellingen.

Beide diensten kunnen een impuls geven aan de toegankelijkheid van literatuur via internet. Via Google Books kan een enorme hoeveelheid oudere publicaties online toegankelijk worden gemaakt. Google Scholar vormt bovendien (samen met andere zoekmachines) een steun in de rug voor vrij beschikbare literatuur via *repositories*: wetenschappers krijgen door de combinatie van elektronische opslag en effectieve zoekmachines een nieuwe etalage voor vrij toegankelijke literatuur. (Zie ook Open Access en Digital Academic Repositories).

### **Amazon**

De internetboekhandel Amazon heeft een vergelijkbare dienst als Google Books: 'Search inside this book'. Maar Amazon wil deze dienst verder uitbouwen met 'Amazon Pages', waarmee lezers online toegang krijgen tot bepaalde pagina's, delen of hoofdstukken van een boek. Wetenschappers kunnen op die manier bijvoorbeeld enkel de voor een onderzoek relevante stukken uit publicaties kopen. Daarnaast is 'Amazon Upgrade' aangekondigd, waarmee kopers van een boek ook toegang krijgen tot de elektronische versie.

### **Open Content Alliance**

Dit initiatief kijkt op een aantal belangrijke punten af van het programma van Google. Het programma richt zich meer op rechtenvrije boeken (*public domain*) en stelt de inhoud hiervan beschikbaar aan ieder die hier gebruik van wil maken, bijvoorbeeld aan andere zoekmachines. Het programma staat ook open voor andere werken, zoals film en muziek en het is in beginsel niet gratis voor de aanbieders, zoals Google Books.

In de Open Content Alliance (<http://www.opencontentalliance.org/>) werken diverse partijen samen: Yahoo! indexeert het materiaal en financiert de digitalisering van de door de Universiteit van Californië geselecteerde literatuurcollectie; Internet Archive beheert het systeem waarop de digitale werken beschikbaar worden gesteld; Adobe en HP leveren de technologie voor digitalisering; en diverse andere partijen, zoals de Universiteit van Toronto en uitgeverij O'Reilly leveren boeken.

Als onderdeel van de Open Content Alliance werkt ook MSN aan de digitalisering van boeken. MSN begint met ongeveer 100.000 boeken van de British Library en zal in de loop van 2006 een eigen zoekmachine voor boeken introduceren. *MSN Book Search* zal vermoedelijk net als Amazon ook aanvullende diensten leveren, zoals inzage in en verkoop van delen van boeken.

Literatuur: Andreas von Bubnoff, Science in the web age: The real death of print, Nature 438 (7068), 550-2 (01 Dec 2005)

Terug naar artikel