

Open Library: biblioteca digitale 2.0

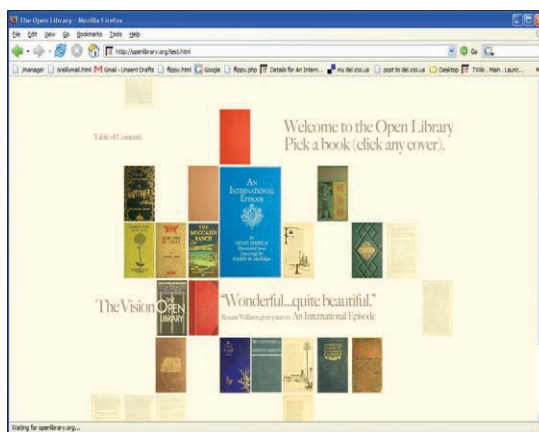
Fabio Di Giammarco

Stimolante e al tempo stesso controverso, fa discutere l'ultimo prodotto della fucina dell'Internet Archive

*Biblioteca di Storia moderna e contemporanea
Roma
digiammarc@tiscali.it*

“Nostro scopo è raccogliere ogni libro esistente – in commercio, fuori catalogo, antico, disponibile soltanto in biblioteca, digitalizzato, in solo formato testo ecc. – e creare un libero accesso sul web a questo materiale librario proveniente da tutto il mondo.” A prima vista potrebbe sembrare l'ultima mission stile Google, per annunciare l'avvento dell'ennesima biblioteca universale. Ma questa volta, nel *mare magnum* delle iniziative di biblioteche digitali on line pare invece spuntare qualcosa di nuovo, qualcosa che da più parti viene descritto come rivoluzionario. Si tratta dell'Open Library di Aaron Swartz,¹ ultimissimo prodotto della fucina dell'Internet Archive,² ovvero del più grande e dinamico repository “melting pot” on line del mondo.

Tra “2.0” e aspirazioni borghesiane, il progetto di questa nuova biblioteca digitale gioca le sue carte puntando su un modello non convenzionale, aperto, collaborativo e universale. Il tutto basato su alcune innovazioni tecnologiche e biblioteconomiche che potrebbero segnare una svolta nei progetti di *digital libraries*, ma che intanto fanno discutere nonché storcere il naso a più di un bibliotecario. Roy Tennant, della California Digital Library³ nonché collaboratore dell'OCLC,⁴ in un suo intervento su Library Journal.com ha in pratica



La home page di Open Library

liquidato l'operazione, e non tanto per le velleitarie pretese “universalistiche”, quanto per la poca serietà d'impianto: un catalogo generato e gestito dagli utenti (!) e un nuovo “metadata schema” che, una per tutte, prevede un campo autore, a suo dire, poco affidabile...⁵ Comunque, quello che ad alcuni sembra poco convincente è proprio il “core” della Open Library, vale a dire il suo nuovo modello tecno-organizzativo centrato sul cosiddetto utente “UGC”.⁶ Un sistema ispirato al “web partecipativo” e imperniato su un nuovo formato di catalogazione, un nuovo database e un sistema wiki⁷ per la gestione dei dati e contenuti. FutureLib è la novità per quanto riguarda lo “schema metadati”, uno standard che a differenza del MARC⁸ tiene conto non solo delle esigenze dei bibliotecari ma anche di quelle degli editori (vedi formato ONIX)⁹ e soprattutto degli utenti. Su piatta-

forma PostgreSQL¹⁰ gira invece ThingBD, nuovo modello di database flessibile per l'organizzazione dei dati. E infine c'è Infogami, un software wiki particolarmente duttile, in grado di manipolare e ricombinare i dati potenziando *in primis* le caratteristiche di ThingBD, ma implementando anche tutte le altre articolazioni del sistema (le pagine del website, i tool di catalogazione, i template¹¹ ecc.) con un ruolo – all'interno del modello proposto

– così importante che il progetto “Open Library” potrebbe anche definirsi: “Un wiki globale per catalogare, archiviare e socializzare libri d'ogni epoca e genere”. Tuttavia, l'approccio “2.0” – stimolante e al tempo stesso controverso – realizza solo una parte di quella “biblioteca digitale aperta” immaginata da Aaron Swartz. L'altra è data dall'adesione al movimento “open access”. Il progetto fa, infatti, parte dell'Open Content Alliance¹² – centro di aggregazione di progetti di digitalizzazione con standard condivisi – e di conseguenza adotta il protocollo OAI-PMH,¹³ cioè la tecnologia emergente per disegnare una cooperazione a 360° in rete tra comunità interessate, in questo caso bibliotecari, istituzioni, editori e utenti. Già oggi, il cosiddetto mondo “open access” è una realtà di milioni di documenti digitalizzati (repository OAI-PMH compatibili)

sparsi nel web, e con l'harvesting, tramite il protocollo OAI-PMH, diventa possibile raggiungere collezioni e risorse che interessano, prelevare i metadati e popolare, aggiornare e arricchire i propri OPAC e cataloghi, come quello dell'Open Library. Anche Google, come tutti sanno, sta costruendo, con grande dispiego di mezzi e di risorse finanziarie, una biblioteca digitale universale. Ma la sua incerta posizione rispetto all'open access, più precisamente sulla condivisione degli standard di codifica, potrebbe addirittura compromettere il progetto, creandogli in futuro grossi problemi dal punto di vista dell'interoperabilità e onerosi interventi di conversione.

Riguardo poi alle aspirazioni "universalistiche" dell'Open Library – messo in conto l'inevitabile scetticismo che accompagna questo tipo di "annunci ad effetto" – il progetto, a differenza di altri ambiziosamente lanciati negli ultimi tempi sul web, ma poi spesso attestatesi in una sorta di limbo, può effettivamente, da questo punto di vista, mettere in campo qualcosa in più, cioè il fatto di agire nell'ambito dell'Internet Archive e dell'Open Content Alliance: due grandi catalizzatori di risorse sul web. In particolare, l'OCA ha mostrato di voler perseguire – a cominciare dall'interoperabilità – una politica davvero globale con il coinvolgimento di università, istituzioni culturali e biblioteche di tutto il mondo, così da fornire un modello e basi per sviluppare effettivi progetti di biblioteche universali.¹⁴ Non solo, l'Internet Archive offre già gratuitamente – a parte tutti gli altri formati multimediali stipati nel suo mega-archivio web – oltre 200.000 libri in formato digitale; mentre, ritornando all'Open Content Alliance, va anche aggiunto che la sua – diversamente da quella di Google – può essere ritenuta una "mass digitization di qualità" sia dal punto di

vista dell'acquisizione che da quello dell'accesso e dei servizi.¹⁵ Insomma, due partner per l'Open Library di tutto rispetto, capaci di garantirgli un vasto orizzonte di riferimento.

Viste potenzialità e prospettive, sorge, a questo punto, spontanea un'ultima domanda: al momento, il progetto di Aaron Swartz cosa offre? Non poco. Il suo debutto sul web è progredito attraverso la pubblicazione di due siti. Il primo,¹⁶ creato direttamente dall'Internet Archive, è un classico "sito vetrina", ricco di informazioni sul progetto ma realizzato soprattutto per mostrare come i libri possono essere suggestivamente rappresentati on line. Il secondo, assai più interessante, è invece un vero e proprio "demo-site".¹⁷ Scorrendo quest'ultimo, ci si può fare un'idea abbastanza precisa su quello che è, che sta diventando e che sarà l'Open Library. Utilizzando, ad esempio, l'opzione "Guided tour" si possono testare le diverse interfacce con le possibili risposte del catalogo on line rispetto alle principali tipologie di risorse trattate: libri in commercio, libri fuori catalogo e/o antichi già acquisiti in formato digitale, oppure altri generi di documenti, a cominciare dagli articoli scientifici. Ovvio discriminante anche nel caso dell'Open Library, come del resto in tutti i progetti di *digital libraries*: l'accesso ai contenuti, vale a dire la questione copyright. Anche in questo caso la disponibilità "full text", con eventuale download e scelta tra diversi formati,¹⁸ riguarda solo i volumi di pubblico dominio, mentre per i libri in commercio – più o meno recenti – coperti da diritto d'autore vengono solitamente offerti un estratto (il più delle volte il primo capitolo), cenni biografici dell'autore, link al suo website e alle principali librerie on line della rete qualora si voglia procedere all'acquisto di una

copia. Infine, a ribadire la carica innovativa del progetto anche e soprattutto dal punto di vista della comunicazione, c'è il blog "Disruptive Library Technology Jester" che propone "l'Open Library demonstration screencast":¹⁹ un video di circa dodici minuti per lasciarsi guidare tra le pagine del demo-site con tanto di dimostrazioni e spiegazioni, una carrellata per gettare un inedito colpo d'occhio su un progetto di biblioteca digitale.

Note

¹ <<http://www.aaronsw.com/weblog/>>.

² <<http://www.archive.org/index.php>>.

³ <<http://www.cdlib.org/>>.

⁴ <<http://www.oclc.org/>>.

⁵ <<http://www.libraryjournal.com/blog/1090000309/post/1800011980.html>>.

⁶ User-Generated Content.

⁷ In questo caso il termine "wiki" si riferisce a un software collaborativo utilizzato per creare siti web.

⁸ Acronimo di "Machine Readable Cataloguing" (Catalogazione leggibile da una macchina), una specifica per la rappresentazione dell'informazione bibliografica.

⁹ È il formato standard internazionale per la rappresentazione e la comunicazione di informazioni bibliografiche in formato elettronico all'interno dell'industria del libro.

¹⁰ È un completo database relazionale ad oggetti con licenza libera.

¹¹ È una pagina speciale creata per essere inclusa in altre pagine.

¹² <<http://www.opencontentalliance.org/>>.

¹³ Protocol for Metadata Harvesting.

¹⁴ Il progetto ha un'impostazione prettamente bibliotecaria con la partecipazione di decine e decine di istituzioni bibliotecarie.

¹⁵ Digitalizzazione di qualità con il sistema SCRIBE, formato di archiviazione PDF, servizi di print on demand.

¹⁶ <<http://www.openlibrary.org/>>.

¹⁷ <<http://demo.openlibrary.org/>>.

¹⁸ Tra gli altri il formato DjVu: una tecnologia innovativa ad alta compressione per visualizzare testi digitalizzati.

¹⁹ <<http://dl.tj.org/2007/07/open-library/>>.