

Il Progetto Google books: *la prima grande esperienza di accesso diretto al patrimonio bibliografico nazionale*

Osvaldo Avallone

Direttore Biblioteca Nazionale Centrale "Vittorio Emanuele II" - Roma

A seguito dell'Accordo siglato tra Mibac e Google, nel novembre 2012 ha preso ufficialmente avvio il Progetto Google Books che vede coinvolte, oltre alla due Biblioteche Nazionali Centrali di Roma e di Firenze, anche la Biblioteca Nazionale di Napoli e l'Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche. Obiettivo del progetto è quello di digitalizzare circa 340.000 volumi e di metterli in rete sia attraverso Google, sia attraverso la rete delle Biblioteche Pubbliche Italiane e da altri siti istituzionali della Pubblica Amministrazione.

Nel contributo, oltre a descrivere gli aspetti scientifici del Progetto (i criteri di selezione delle opere da digitalizzare, le problematiche relative al diritto d'autore) e gli aspetti di gestione amministrativa (la Biblioteca Nazionale Centrale di Roma è coordinatrice esecutiva dell'intero progetto), si riflette su alcuni aspetti critici e questioni non risolte che spesso hanno contraddistinto i numerosi progetti di digitalizzazione nati in questi anni.

Da quando si è iniziato a parlare di digitalizzazione in Italia, abbiamo assistito ad un fiorire di progetti di ogni genere e tipo, che avevano in comune tra loro la finalità di rendere disponibile l'accesso virtuale diretto agli oggetti digitalizzati. E qui finiscono le analogie, perché quanto a standard, metodologie, filosofie di scelta e obiettivi ultimi si era ben lungi dall'omogeneità. Data l'occasione presente e la mia funzione istituzionale, mi limiterò ad occuparmi del patrimonio bibliografico, la cui complessità di gestione non è inferiore a qualsiasi altro oggetto che si pensi di sottoporre a digitalizzazione. La possibilità e forse perfino l'esigenza di disporre di copie digitali delle unità bibliografiche non è proprio recentissima e scaturisce dai microfilms, cioè quella che era stata fino a circa 15 anni fa l'unica possibile risposta all'esigenza da cui nasce il digitale, che si può riassumere in tre diverse motivazioni:

- a. Rendere consultabile da remoto;
- b. Tutelare l'integrità fisica dell'unità bibliografica dal degrado da uso;
- c. Ampliare la fascia di utilizzo.

Va da sè che tali finalità comportavano, quasi di default, anche una serie di vantaggi economici, che rendevano ancora più appetibile l'investimento necessario per digitalizzare. La prima esigenza che si avvertì però fu quella, giustamente, di programmare "piani di digitalizzazione", per non disperdere in iniziative a "macchia di leopardo" i vari progetti che cominciavano a vedere la luce. Siamo a circa dieci anni fa, c'era già la crisi o una delle tante crisi che ci hanno afflitto, ma ancora qualche risorsa era disponibile. I momenti essenziali furono quindi individuati in: scelta delle tipologie di materiale da digitalizzare, selezione delle unità bibliografiche, individuazione dei motori di ricerca (o delle funzionalità di sistema) che non riducessero i nascenti data-bases a delle semplici (ed inutili) successioni d'immagini. Personalmente sono dell'opinione che per la buona riuscita di operazioni del genere è necessario che i bibliotecari dimentichino se stessi, agiscano cioè come se non fossero bibliotecari, ma semplici utenti.

Questa mia posizione, più volte espressa e più volte ovviamente contestata, si è rafforzata ulteriormente di fronte ai risultati, malgrado le buone premesse, che derivarono dai piani di digitalizzazione nazionale.

Per carità di patria non entro nei dettagli, ma basterà considerare che l'effetto finale fu esattamente l'opposto di quello che si cercava in premessa: ognuno digitalizzò quello che riteneva più importante o quello che gli faceva più comodo: chi digitalizzò giornali, riviste e ritagli di giornali, chi i cataloghi antichi, chi i manoscritti, chi addirittura collane o enciclopedie di cui non c'era più nessun bisogno e che non si potevano certo considerare rare. Il tutto, aggravato spesso da inutili duplicazioni, provocò l'inopinato effetto che di leopardi ormai ce n'era una mandria, ed era avvenuta una cospicua dispersione di risorse senza risultati di livello adeguato.

Non furono neanche errori da cui imparare, perché quello che dovevamo imparare lo sapevamo già, ma non avevamo potuto metterlo in pratica. L'unica cosa che apparve ben chiara fu che la digitalizzazione era da considerare un investimento, costoso, che andava rigidamente organizzato e monitorato "step by step", come va di moda dire ora, oppure passo passo com'è chiarissimo anche in Italiano! In un clima che vedeva decurtate in maniera intollerabile le risorse destinate alla cultura, avvenne il primo contatto con Google, circa 7 od 8 anni fa. Fu un contatto breve in cui i rappresentanti di Google ed i vertici dell'Amministrazione evidentemente non si intesero, dato che non sortì risultato alcuno. Solo nel 2010 l'approccio con Google venne cercato e recepito in maniera da giungere a risultati concreti. In breve Google si offrì di digitalizzare il patrimonio bibliografico posseduto dalle due Biblioteche Nazionali Centrali di Roma e Firenze a "costo zero" (musica per le orecchie ministeriali e politiche) e iniziarono le trattative alle quali parteciparono il Direttore Generale delle Biblioteche, il Direttore Generale per la valorizzazione del patrimonio, i due Direttori delle due Biblioteche Nazionali Centrali ed i rappresentanti di Google, numerosi e

quasi mai gli stessi. C'erano una serie di problemi su cui fare chiarezza: Google non poteva certo pensare ad un programma così ambizioso da digitalizzare quasi 14 milioni di unità bibliografiche! Si imponeva un tetto e una scelta. In più si aggiunsero in itinere, come partners del progetto, la Biblioteca Nazionale di Napoli, una delle più grandi e ricche d'Italia e l'Istituto Centrale per il Catalogo Unico delle Biblioteche italiane per la parte di standardizzazione dei dati e per gli adeguamenti a carattere informatico e tecnologico che si fossero resi necessari. Emerse anche che il famoso "costo zero" era una semplice astrazione, visto che le Biblioteche, depauperate di risorse e più ancora di personale, non erano in grado di fare fronte con i loro mezzi alle esigenze materiali di un progetto ad elevato tasso di complessità. Vennero reperiti i fondi presso la Presidenza del Consiglio dei Ministri, poi decurtati, ma pur sempre sufficienti a realizzare una prima tranche di progetto. Venne redatto il progetto esecutivo a cura della Biblioteca Nazionale di Roma che era stata individuata come coordinatrice esecutiva dell'intero progetto e a cui faceva capo la gestione amministrativa di tutti i partners. Non è questa la sede per dilungarsi sulle difficoltà tecniche dell'impresa che, posso giurare, sono state numerosissime quantitativamente ed estremamente laboriose sotto l'aspetto tecnico-scientifico. Basti ricordare, a solo titolo esemplificativo, che le bozze di progetto esecutivo sono state circa 12 prima di approdare a quella finale, che poi "finale" non è mai e cercherò di spiegarne in seguito il perché. Il tutto senza ricorrere a nessun apporto esterno, come consulenze tecniche e simili. Intanto Google aveva esplicitato una serie di preoccupazioni da prendere in considerazione come limiti oggettivi di progetto. Innanzitutto, per evitare noie con la legislazione sul diritto d'autore, bisognava limitarsi alle opere pubblicate entro il 1900 (poi, per essere più realisti del re, si scese al 1870); inoltre, per motivi di compatibilità con le modalità operative di Google, andavano escluse dalla digitalizzazione le opere che avessero al loro interno tavole ripiegate (il che con i libri del 600/700 è molto frequente), le opere di formato "fuori standard", le opere in condizioni di conservazione non ottimali e soprattutto, per motivi assicurativi, le opere che superassero la soglia di valore di 10.000 euro. Il che tagliava la testa alle speranzose dichiarazioni di alcuni quotidiani che avevano magnificato la possibilità di vedere in digitale la prima edizione del Principe di Machiavelli o del Sidereus Nuncius di Galilei. In più, limitando la digitalizzazione ad opere del 600, 700 e 800 (non tutto), rimaneva automaticamente escluso il materiale più prezioso, dagli incunaboli alle cinquecentine. Tali compressioni, insieme con la riduzione di risorse, comportavano che il primo modulo avrebbe prodotto la digitalizzazione di 340.000 volumi. Il che è pur sempre un bel risultato, ma ancora lontano dal limite previsto nell'accordo originario (1.000.000 di volumi). Veniva superata anche un'ulteriore difficoltà, sorta dal fatto che Google intendeva inviare i volumi da digitalizzare ad uno scanning center in Germania. L'opposizione dei Direttori e la ferma presa di po-

sizione del Segretario Generale, spingeva Google ad installare uno Scanning Center in Italia, con grande soddisfazione di tutti.

Predisposto quindi il progetto esecutivo, istituito il Comitato Scientifico per la verifica dei risultati e dei rendiconti, avviate tutte le complesse attività preliminari, coordinati i cronoprogrammi di tutti gli Istituti coinvolti, acquisite le tecnologie e gli adeguamenti tecnologici necessari, la macchina da guerra si è messa in moto, partendo agli inizi di Dicembre 2012 dalla Biblioteca Nazionale Centrale di Roma, che inviava i primi venti carrelli carichi di volumi con molta apprensione, ma anche con entusiasmo e speranza. Seguiva a Febbraio 2013, con puntualità cronometrica la Biblioteca Nazionale Centrale di Firenze, ad Aprile la Biblioteca Nazionale "Vittorio Emanuele III" di Napoli.

Il progetto esecutivo in questi casi è un "work in progress", dal momento che fino a che non si va "in operativo" non si sa cosa si trova realmente; le stime sono fatte obbligatoriamente con approssimazione ed è necessario, di volta in volta, adeguare e "tarare" nuovamente il progetto, che si rileva in fine come "traccia", dettagliatissima, ma flessibile per forza di cose.

Il progetto comunque è partito rispettando le previsioni e sta proseguendo, con i dovuti scongiuri, con pochissime difficoltà. Volendo tracciare un bilancio ad oggi, possiamo dire che sono stati inviati allo scanning center di Google 176 carrelli per un totale complessivo tra le tre biblioteche di 45.760 volumi. Sono stati restituiti digitalizzati oltre 30.000 volumi. Sono stati impiegati con regolari contratti a tempo determinato n. 51 giovani per le esigenze di movimentazione dei volumi, di catalogazione, di verifica dello stato di conservazione ed eventuali interventi di restauro e per la gestione amministrativa del progetto.

Una domanda che mi sento spesso porre è: "tutte luci e niente ombre?" e "ma Google cosa ci guadagna a digitalizzare gratis, ad installare e pagare uno scanning center ecc.? ". La mia risposta è che seppure ci fossero delle ombre, le luci sono superiori ad esse. Senza Google questo progetto non si sarebbe mai realizzato, questo è certo. L'Amministrazione non dispone, e con l'attuale situazione non disporrà mai, delle risorse necessarie per un'operazione di tale portata. Google non ha preteso l'esclusiva se non nei confronti dei concorrenti. Le digitalizzazioni verranno messe in rete da Google, ma anche dalla rete delle Biblioteche Pubbliche Italiane e da altri siti istituzionali della Pubblica Amministrazione. Certo non se ne potrà fare un uso commerciale, ma questo le Biblioteche già non lo fanno. In ogni caso si è evitato di subordinarsi ad altri accordi, come in passato, in cui è stata operata una vera e propria cessione di diritti in cambio della digitalizzazione! E dopo questo? La realtà attuale rimane sempre "leopardata": chi si crea basi dati digitali tematiche sui soggetti più bizzarri, chi dà libero sfogo ad ambizioni personali; da ogni lato viene chiesta collaborazione e disponibilità, ma le Biblioteche avrebbero tantissimo da dire e da fare in proprio.

Occorrono solo risorse, coordinamento, professionalità: la luna nel pozzo! Google rappresenta un'opportunità, e anche una speranza, se questo nostro Paese saprà ricordare che uno dei suoi motori trainanti forse il più potente è quello dell'arte e della cultura, della tradizione e dell'identità culturale che trova le sue radici proprio nella memoria storica custodita dalle Biblioteche italiane.

Following the Agreement signed by Mibac and Google, in November 2012, the Project Google Books officially started up. The Libraries involved are the two National Central Libraries of Rome and Florence, as well as the National Library of Naples and the ICCU (the Central Institute for the Union Catalogue of Italian libraries and for bibliographic information). The aim of the Project is to digitize more or less 340.000 books and put them online both through Google and through the Italian Public Libraries network and other institutional sites of the Italian Public Administration.

In the contribution, besides the description of the scientific aspects of the Project (the selection criteria of the books to digitize, the problems related to the copyright) and of the administrative management ones (the National Central Library of Rome is the executive coordinator of the whole Project), there is a careful consideration on some critical aspects and unsolved problems that have often characterized the several digitization Projects that have been starting up these last years.