

Il catalogo delle biblioteche dell'Università di Genova in Worldcat

Antonio Scolari

Nel mese di ottobre è stato completato il riversamento dei dati del catalogo delle biblioteche dell'Ateneo su WorldCat (<<http://www.worldcat.org/>>), l'OPAC di accesso ai dati del maggiore catalogo cooperativo internazionale, a cui partecipano biblioteche di tutto il mondo, gestito da OCLC. Il catalogo cooperativo di OCLC contiene attualmente oltre 273 milioni di record catalografici a cui corrispondono circa 2 miliardi di copie possedute da 72.000 biblioteche di svariati paesi. Partecipano al catalogo le principali biblioteche di ricerca delle università statunitensi ed europee, 44 biblioteche nazionali e importanti biblioteche pubbliche. Il catalogo iniziato nel 1971 come uno strumento di cooperazione statunitense, nel corso degli anni ha assunto un carattere sempre più internazionale: oggi nel catalogo sono comprese pubblicazioni di 470 lingue diverse, mentre oltre il 60% dei record catalogati appartiene a lingue diverse dall'inglese¹.

L'iniziativa, attuata dal Centro di Servizi per il coordinamento del Sistema Bibliotecario di Ateneo, fa parte di un progetto del Consorzio di Sistemi Bibliotecari di Ateneo CIPE (<<http://www.unicipe.it>>), di cui l'Università di Genova è membro fondatore. Attualmente sono stati caricati nel catalogo OCLC i record bibliografici delle università di Genova e Siena e nei prossimi mesi si aggiungeranno quelli degli altri nove atenei membri del consorzio CIPE (Bologna, Firenze, Modena e Reggio, Padova, Parma, Pisa, Politecnica delle Marche, Sassari, Venezia - Cà Foscari). Si prevede che l'insieme dei dati catalografici degli undici atenei del Consorzio arricchisca il catalogo internazionale di oltre 10 milioni di localizzazioni di libri, offrendo quindi una significativa rappresentazione della ricchezza del materiale bibliografico posseduto dalle biblioteche universitarie del nostro paese.

La partecipazione a WordCat è una delle iniziative di apertura e condivisione del catalogo perseguite dal Sistema Bibliotecario di Ateneo e va ad affiancarsi alla recente apertura del Polo SBN "SGE", in collaborazione con il Sistema bibliotecario urbano.

Una prima ricaduta riguarda il catalogo dell'Ateneo: pur se strutturato in Unimarc, al momento di estrarre le registrazioni per il loro riutilizzo in un altro catalogo, si è resa necessaria una revisione assai estesa soprattutto per quanto riguarda aspetti che, lavorando a livello locale, vengono normalmente sottovalutati, ma che in un catalogo cooperativo, specie se di tipo

¹ I dati sono tratti da *OCLC Annual Report 2011-2012*, il più recente dei report annuali, disponibile all'indirizzo

<www.oclc.org/content/dam/oclc/publications/AnnualReports/2012/2012.pdf> e dalla pagina *WorldCat facts and statistics* <<http://www.oclc.org/en-europe/worldcat/catalog.html>>

internazionale, sono considerati come elementi di base per il caricamento dei record derivati da cataloghi esterni. Inoltre il catalogo dell'Ateneo nella sua parte meno recente è frutto di un passaggio, effettuato alla fine degli anni '90, dal vecchio sistema Aleph che utilizzava un formato proprietario dei record bibliografici all'attuale Aleph500 che, per le installazioni italiane, utilizza il formato Unimarc. Ciò ha comportato nel corso del progetto una speciale revisione dedicata a questi record, che in molti casi presentavano un tracciato MARC non completamente conforme.

A seguito dell'invio ad OCLC nel novembre 2012 di un primo lotto di prova di circa 80.000 record in formato ISO 2709, sono state riscontrate numerose problematiche ricorrenti nei record di campionatura inviati, in particolare in dati ritenuti da OCLC obbligatori al fine del caricamento dei record. Così, ad esempio, è stato necessario intervenire sull'etichetta 100 (Dati di elaborazione) per aumentare la coerenza tra il tipo di data indicato in posizione 8 e le date registrate nelle seguenti posizioni; nell'etichetta 101 (Lingua della pubblicazione) figuravano codici di lingua non corretti rispetto agli standard normalmente utilizzati; o ancora un gran numero di record era privo dell'etichetta 102 (Paese di pubblicazione), un dato invece obbligatorio in WorldCat, come del resto anche in SBN. In quest'ultimo caso è stato necessario predisporre una procedura che ha consentito, non senza intervento manuale, di aggiungere o correggere il campo in oltre 300.000 record del catalogo.

Simile, ma affrontato in modo differente, il caso del campo 801 (Fonte di provenienza del record), che nel nostro catalogo non presenta contenuti costantemente omogenei, in particolare per quanto concerne l'indicazione dell'ente responsabile della creazione o modifica del record. Per rispondere in modo adeguato alla richiesta di OCLC di una totale uniformità di contenuti del campo in tutti i record inviati, si è deciso di richiedere la codifica dell'Ateneo nella *MARC Code List for Organizations*, la lista internazionale delle biblioteche e dei sistemi bibliotecari gestita dalla Library of Congress², che rilascia i codici alle istituzioni non solo statunitensi, per l'uso all'interno dei record MARC per progetti di cooperazione nazionali e internazionali. L'Ateneo è quindi stato codificato con la sigla "IT-GeUSG" (in forma normalizzata: "itgeusg") e tale sigla è stata inserita nel sottocampo \$b dei record esportati. In questo caso non si è intervenuti direttamente nei record, ma si è impostata una procedura di esportazione che sostituisce i campi 801 dei record e ne crea uno specifico ed identico per tutti i record solo al momento della creazione dei file da inviare a OCLC, senza operare modifiche nel database locale.

Il miglioramento del catalogo non ha riguardato solo questi aspetti in parte "formali" del MARC, ma su segnalazione di OCLC si è intervenuti anche su problemi di maggiore importanza dal punto di vista della correttezza dei dati. Così, ad esempio, sono state riviste alcune centinaia di titoli erroneamente

² Si veda: <<http://www.loc.gov/marc/organizations/>>.

privi del sottocampo \$a e quindi introvabili nel catalogo; sono stati corretti i delimitatori di articolo nei campi titolo (200. 500, 454, ecc.) in un alto numero di casi; sono stati sanati in svariati campi sottocampi erronei; si è intervenuti in moltissimi record descritti a più livelli per i quali non erano coerenti i dati di legame; si sono arricchiti record bibliografici – soprattutto di opuscoli o estratti – che risultavano troppo lacunosi per una loro corretta identificazione. Questi ultimi tipi di revisioni hanno aumentato la coerenza del catalogo e quindi hanno anche portato a un complessivo miglioramento della qualità del catalogo per gli utenti dell'OPAC locale.

Le revisioni sul catalogo sono state effettuate nella primo semestre del 2013, mentre nel contempo venivano predisposte le procedure per l'estrazione dei dati e per effettuare alcune elaborazioni sui dati al momento dell'estrazione. È il caso, ad esempio, della eliminazione di campi proprietari sia di tipi numerico che di tipo alfabetico, utilizzati per diversi scopi nel catalogo, ma che non dovevano essere trasmessi nell'esportazione finale. È stato possibile predisporre queste procedure grazie alla disponibilità in Aleph di un linguaggio interno che permette la creazione di *routine* da applicare ai record bibliografici e manipolarne in modo adeguato il contenuto, senza l'intervento da parte dell'azienda che produce il software.

Sono stati esclusi dall'esportazione tutti i record che erano privi di copia, per l'impossibilità da parte di OCLC di localizzarli è il caso degli spogli, di risorse elettroniche remote e di record appartenenti a catalogazioni a più livelli; sono stati inoltre escluse pubblicazioni che risultavano in acquisizione al momento dell'estrazione finale.

Nel mese di luglio sono stati depositati sul server olandese di OCLC i dati esportati, in formato ISO2709, nella loro versione finale; i file sono stati processati dalla sede americana nel corso dell'estate, per giungere – non senza qualche ritardo – al loro definitivo caricamento in WorldCat nel corso del mese di ottobre.

OCLC ha processato 559.437 registrazioni bibliografiche, che rappresentavano, alla data dell'estrazione, i record bibliografici con almeno una copia appartenente a una delle biblioteche dell'Ateneo; il totale dei record bibliografici dotati di copie in catalogo ammontava, alla medesima data, a 802.545, conteggiando anche quelli che riportavano almeno una copia appartenente a una delle biblioteche civiche. Sono stati rifiutati dalle procedure di importazione 833 record, che contenevano errori non sanabili. Infine sono state create 667.229 copie riferite alle biblioteche dell'Ateneo. Sono stati creati in WorldCat 365.465 nuovi record: l'alto numero di nuove creazioni si spiega in parte con il fatto che finora nel catalogo non sono moltissimi i record con lingua di pubblicazione "ita" (circa 6,3 milioni prima del caricamento di Genova) e che sono relativamente scarsi i record con lingua di catalogazione "ita" (circa 1.4 milioni prima del caricamento di Genova). Tuttavia l'impressione è che la deduplicazione automatica, pur basata su algoritmi avanzati, sia ancora non completamente efficace e che quindi un certo numero

di record «nuovi» rappresentino in realtà delle duplicazioni di registrazioni già esistenti nel catalogo.

A valle dell'importazione OCLC ha rilasciato all'ateneo l'elenco degli "OCL numbers", gli identificativi univoci, attribuiti ai record di Genova in WorldCat, che sono stati importati nel catalogo nel campo 035, in modo che sia possibile identificare in modo univoco i record del nostro catalogo in quello cooperativo. Sarà così possibile procedere a future manutenzioni e, soprattutto, tra breve verrà creato un link diretto fra i record WorldCat e il nostro Opac locale a cui verranno reindirizzati gli utenti di WorldCat.

Il nuovi record via via catalogati dalle biblioteche dell'ateneo verranno inviati a OCLC con cadenza trimestrale, per garantire un ragionevole aggiornamento dei dati nel catalogo internazionale.

Fin qui i risultati e le ricadute per così dire «interne» del progetto, che però, in considerazione della sua valenza consortile, ha ricadute «esterne» di non minore importanza.

Una prima ricaduta è la visibilità del posseduto di un significativo numero di biblioteche di università italiane nel maggiore catalogo internazionale, a cui gli utenti accedono non solo direttamente, ma anche tramite il collegamento esistente fra Google Books e WorldCat. Nelle poche settimane intercorse dal caricamento dei record sono già giunte a nostre biblioteche richieste di document delivery da biblioteche europee e statunitensi, a riprova della vasta diffusione di WorldCat come strumento per il prestito interbibliotecario.

Più in generale il progetto del consorzio ha lo scopo di colmare una grave lacuna di rappresentazione del posseduto dalle biblioteche italiane nel catalogo di OCLC. Infatti finora la presenza di biblioteche italiane nel catalogo internazionale era assai limitato, al punto che un utente italiano di WorldCat trovava la disponibilità di libri editi in Italia o di importanti autori italiani non in biblioteche italiane, ma soltanto in biblioteche straniere.

Infine è recentissimo l'annuncio di un accordo tra OCLC e Plum Analytics per avviare l'utilizzo di WorldCat per il calcolo di indici bibliometrici per la valutazione dell'impatto dei lavori scientifici dei ricercatori, tramite il prestito interbibliotecario. La notizia è assai interessante, perché potrebbe diventare una via per valorizzare nelle analisi bibliometriche applicate ai prodotti della ricerca il peso delle monografie, un aspetto della massima importanza per i ricercatori delle aree disciplinari umanistiche e sociali, spesso non coperte dai sistemi di calcolo bibliometrico tradizionali, limitati come sono ai periodici. La presenza nel catalogo internazionale di una parte rappresentativa delle pubblicazioni scientifiche in lingua italiana o edita in Italia è di sicuro un veicolo importante per aumentarne la diffusione e l'uso.

Il progetto di CIPE ha poi favorito la traduzione in italiano dell'interfaccia di consultazione di WorldCat, che da qualche mese è disponibile per gli utenti dall'Italia e si affianca alle numerose versioni in varie lingue in cui l'interfaccia è stata tradotta.

vedi anche

Notiziario della Sezione Ligure dell'Associazione Italiana Biblioteche

Numero 2 Vol. 23 Anno 2013 – ISSN 2281-0617

A seguito dell'iniziativa del Consorzio CIPE, anche altri atenei italiani stanno per riversare i propri dati in WordCat, fra questi l'Università Cattolica.



The screenshot displays the OCLC WorldCat interface. At the top, there is a navigation bar with 'Home' and 'Cerca' (Search) menus. Below this is a search bar with a search button and a link to 'Trova una biblioteca' (Find a library). The main content area shows search results for the book 'Atti della giornata in ricordo di Enrico Marchi : Prima Marchi L'. The book details include the author 'Enrico Marchi', the publisher 'Genova : Genova University Press, 2011.', and the format 'Libro : Italian'. The book is listed in the 'WorldCat' database. Below the book details, there is a section titled 'Trova una copia in biblioteca' (Find a copy in a library). This section includes a search bar for the user's location (set to 'genova') and a dropdown menu for the library location (set to '16121 Genova-121, IT'). A 'Trova biblioteche' (Find libraries) button is also present. The search results show one library: 'Università di Genova - Biblioteca di Ingegneria. Primo Polo (CIPE) / University of Genoa - Engineering Library. Section 1 (CIPE)'. The library is located in Genova, Italy. The format of the book is listed as 'Book'.