

CIP e metadattazione condivisa nell'ambiente digitale come nuove forme di cooperazione tra editori e biblioteche

*Laura Manzoni**

Abstract

La diffusione delle risorse digitali e l'informatizzazione delle procedure catalografiche ha determinato la comparsa nel mondo delle biblioteche di soggetti terzi, quali gli editori e i fornitori di risorse elettroniche, che sono i primi a produrre i metadati relativi alle loro pubblicazioni. Da tempo le biblioteche mostrano attenzione verso queste informazioni per impostare le proprie politiche d'acquisto, tuttavia recentemente hanno iniziato ad apparire sempre più chiari i vantaggi che potrebbero derivare da una più stretta collaborazione anche in ambito catalografico. L'aumento esponenziale delle pubblicazioni digitali e la velocità dei cambiamenti dell'universo online (cambi di piattaforme, obsolescenza delle URL, crescita esponenziale dei contenuti open access) sta rendendo impossibile il censimento immediato di tutto ciò che entra a far parte di una biblioteca e la manutenzione che richiedono i metadati relativi alle pubblicazioni digitali. Per questa ragione si sta rivelando necessario, e non più semplicemente auspicabile, definire nuove modalità di cooperazione tra editori e biblioteche. La rilevanza di questo tema è stata recentemente sottolineata dall'IFLA all'interno del documento *Pratiche condivise per le bibliografie nazionali nell'era digitale*, approvato nell'aprile del 2022, all'interno del quale viene dedicata un'intera sezione alla "Cooperazione con editori e produttori di media", i quali vengono considerati una fonte d'informazione primaria per le bibliografie nazionali. In particolare, viene sottolineata l'importanza dei programmi Cataloguing in Publication (CIP) che costituiscono un supporto fondamentale sia per gli editori che per le biblioteche. Esempi significativi di tale collaborazione a livello mondiale sono costituiti dalla Library and Archives Canada che ha sviluppato un programma insieme alla Bibliothèque et Archives nationales du Québec incentrato sulla collaborazione con gli editori di lingua francese con sede in Québec; dalla Library of Congress e dalla British Library. Questi programmi pur

* Ringrazio il prof. Mauro Guerrini per il confronto durante la stesura del saggio e per le successive riletture.

essendo diversi tra loro, poiché cercano di soddisfare delle esigenze locali, si caratterizzano per la stretta collaborazione che si stabilisce tra le agenzie bibliografiche nazionali e gli editori con cui concordano modelli di metadati sulla base dei quali sono costruiti i record inseriti nelle bibliografie nazionali.

Accanto allo sviluppo di queste forme di cooperazione, un altro modo per garantire il pieno riutilizzo dei metadati editoriali da parte delle biblioteche è quello di favorire la mappatura o l'elaborazione di standard comuni tra editori e biblioteche, contenenti tutti gli elementi che essi ritengono fondamentali per perseguire i loro scopi. Un impegno significativo in questo senso è stato assunto dalla Library of Congress che subito dopo la pubblicazione del formato ONIX, adottato in ambito editoriale dal 2000, ha avviato una mappatura con il MARC21. Significativa in questo senso è stata anche la volontà di garantire l'interoperabilità tra ONIX e RDA, *Resource Description and Access*, il primo standard internazionale di metadattazione.

The spread of digital resources and the computerisation of cataloguing procedures has led to the appearance in the library world of third parties, such as publishers and providers of electronic resources, who are the first to produce the metadata relating to their publications. For a long time libraries have been paying attention to this information in order to set up their purchasing policies, but recently the advantages of a closer collaboration also in the field of cataloguing have started to become increasingly clear. The exponential increase of digital publications and the speed of changes in the online universe (changes of platforms, obsolescence of URLs, exponential growth of open access content) is making it impossible to take an immediate census of everything that enters a library and the maintenance that metadata relating to digital publications require. For this reason it is proving necessary, and no longer merely desirable, to define new ways of cooperation between publishers and libraries. The relevance of this issue has recently been emphasised by IFLA in the document *Common Practices for National Bibliographies in the Digital Age* approved in April 2022, in which an entire section is devoted to 'Cooperation with Publishers and Metadata Producers', who are considered a primary source of information for national bibliographies. In particular, the importance of the Cataloguing in Publication (CIP) programmes is emphasised as a fundamental support for both publishers and libraries. Significant examples of such collaboration worldwide are the Library and Archives Canada, which has developed a programme together with the Bibliothèque et Archives nationales du Québec focusing on collaboration with French-language publishers based in Québec; the Library of Congress and the British Library. Although these programmes are different

in that they seek to meet local needs, they are characterised by the close collaboration established between the national bibliographic agencies and the publishers with whom they agree on metadata models on the basis of which the records included in the national bibliographies are constructed. Next to the development of these forms of cooperation, another way to ensure the full re-use of publishing metadata by libraries is to foster the mapping or elaboration of common standards between publishers and libraries, containing all elements they consider essential to pursue their purposes. A significant commitment in this sense was made by the Library of Congress, which immediately after the publication of the ONIX format, adopted in publishing since 2000, started mapping with MARC21. Also significant in this sense was the desire to ensure interoperability between ONIX and RDA, Resource Description and Access, the first international metadata standard.

Keywords: editoria e metadattazione, cooperazione tra editori e biblioteche, publishing and metadata, cooperation between publishers and libraries, Onix, MARC21, RDA, CIP.

L'informatizzazione delle procedure catalografiche ha determinato la comparsa nel mondo delle biblioteche di altri soggetti, quali gli editori e i fornitori di risorse digitali, che sono i primi a produrre i metadati relativi alle loro pubblicazioni. Da un paio di decenni almeno le biblioteche mostrano attenzione verso queste informazioni per impostare le proprie politiche d'acquisto. Tuttavia, negli ultimi anni hanno iniziato ad apparire sempre più chiari i vantaggi che potrebbero derivare da una più stretta collaborazione anche in ambito catalografico. L'aumento esponenziale delle risorse digitali e la velocità dei cambiamenti dell'universo online (cambi di piattaforme, obsolescenza delle URL, crescita esponenziale dei contenuti open access) sta rendendo impossibile il censimento di tutto ciò che entra a far parte di una biblioteca e la manutenzione che richiedono i metadati relativi alle risorse digitali. Per questa ragione si sta rivelando necessario, e non più semplicemente auspicabile, definire nuove modalità di cooperazione tra editori e biblioteche. La rilevanza di questo tema è stata sottolineata dall'IFLA all'interno del documento *Common Practices for National Bibliographies in the Digital Age*¹, approvato nel 2022

1 IFLA, *Common Practices for National Bibliographies in the Digital Age*. 14 luglio 2022, <<https://repository.ifla.org/bitstream/123456789/2001/1/Bibliographies%20Common%20Practices%202022.pdf>>.

e tradotto in italiano con il titolo *Pratiche condivise per le bibliografie nazionali nell'era digitale*². Al suo interno viene attribuita grande rilevanza ai programmi Cataloguing in Publication (CIP), per la condivisione di record bibliografici, a cui si sono affiancate iniziative volte a favorire lo scambio di record d'autorità come l'Author Names Project sostenuto dall'Arcadia Fund dell'Harvard Office of Scholarly Communication e dal Berkman Center for Internet and Society. Fondamentali sono anche alcuni progetti avviati, a partire dall'inizio degli anni Duemila, con l'intento di mappare i formati e gli standard elaborati in ambito biblioteconomico con quelli adottati dagli editori. Infine, particolare attenzione è oggi rivolta alla condivisione di metadati relativi alle risorse digitali, che ormai costituiscono una parte significativa delle collezioni di una biblioteca.

Pratiche condivise per le bibliografie nazionali nell'era digitale

Pratiche condivise per le bibliografie nazionali nell'era digitale dedica un'intera sezione alla *Cooperazione con editori e produttori di media*, i quali vengono considerati una fonte d'informazione primaria per le bibliografie nazionali. In particolare, viene sottolineata l'importanza dei programmi Cataloguing in Publication che costituiscono un supporto fondamentale sia per gli editori sia per le biblioteche. Per gli editori rappresentano un ponte verso il mercato di acquisto delle biblioteche, mentre per quest'ultime costituiscono uno strumento prezioso nella scelta delle risorse da acquisire per arricchire le proprie collezioni. Inoltre, gli accordi CIP consentono alle agenzie bibliografiche nazionali di migliorare la tempestività delle informazioni e, quindi, di aumentare il valore delle bibliografie nazionali per lo sviluppo delle collezioni. Esse stabiliscono quali metadati devono essere forniti dagli editori, mettendo a loro disposizione una scheda informativa stampata o un modulo web da compilare. Gli editori forniscono i dati necessari, con alcuni mesi di anticipo rispetto alla pubblicazione dell'opera, che vengono utilizzati per creare un record preliminare chiamato CIP. Esso verrà poi eventualmente modificato

2 IFLA, *Pratiche condivise per le bibliografie nazionali nell'era digitale*, edizione italiana a cura dell'AIB, Commissione nazionale biblioteche e servizi nazionali. 2023, <<https://repository.ifla.org/bitstream/123456789/2582/1/common-practices-for-national-bibliographies-in-the-digital-age-it.pdf>>.

o arricchito in seguito all'uscita dell'opera e sarà pubblicato sulla bibliografia nazionale. Una copia del record definitivo è, inoltre, fornita all'editore per la stampa sul libro (generalmente sul verso del frontespizio) al fine di facilitarne la catalogazione da parte delle biblioteche che lo acquistano. Nel caso in cui il record CIP non venisse stampato l'editore deve fornire una dichiarazione in cui afferma che il record è disponibile presso l'agenzia bibliografica nazionale.

Non tutte le agenzie bibliografiche nazionali hanno avviato progetti di questo tipo. L'ultima ricognizione condotta dall'IFLA, sulla base di un questionario elaborato dalla Library of Congress risale al 2003/2004 e ha portato alla pubblicazione di un report nel 2005³. Allora risposero al questionario 42 agenzie bibliografiche nazionali di cui 25 avevano programmi CIP attivi. Molti di essi sono sorti in seguito alla pubblicazione del documento *Recommended standards for cataloguing-in publication: the CIP data sheet and the CIP record in the book* emanato dall'IFLA International Program for UBC nel 1986⁴ e incentrato sulla definizione di modelli standard per il modulo di richiesta dei metadati che dovevano essere forniti dagli editori e su come gli editori dovessero stampare la scheda CIP nei libri. Questi programmi condividono una serie di motivazioni di base e di obiettivi quali: favorire la standardizzazione delle registrazioni bibliografiche e lo scambio di metadati, ridurre la duplicazione nella catalogazione, servire i bisogni delle biblioteche e dei lettori, migliorare l'efficienza delle acquisizioni e dei piani di sviluppo delle collezioni, assistere gli editori e i rivenditori di libri, mantenere authority file per gli autori.

Esempi significativi a livello internazionale di programmi oggi molto attivi sono costituiti dalla Library of Congress, dalla Library and Archives Canada e dalla British Library. In generale, queste iniziative si concentrano sulle pubblicazioni monografiche nazionali, sia in formato cartaceo sia elettronico, escludendo tutte le altre tipologie di risorse.

La Library of Congress e la nascita del primo progetto CIP

La Library of Congress (LoC) è stata la prima biblioteca ad avviare un progetto per la catalogazione di libri in corso di pubblicazione «con

³ IFLA, *Survey of CIP Program Report*, <https://www.loc.gov/catdir/cipsurvey/IFLA_CIP_Survey_Report.pdf>.

⁴ IFLA International Programme for UBC, *Recommended standards for cataloguing-in publication: the CIP data sheet and the CIP record in the book*, London: British Library Service, 1986.

l'obiettivo di facilitare e velocizzare la catalogazione e rendere omogenei gli indici di archiviazione e reperimento della registrazione»⁵. L'iniziativa nasce il 1° luglio 1971 con la definizione di un accordo tra la LoC e alcuni editori in base al quale essi sono tenuti a inviare, oltre a una serie di metadati essenziali, anche una copia della risorsa (la versione integrale o parziale)⁶. Per i primi tre decenni gli editori hanno compilato schede cartacee, spedite via posta alla LoC insieme alla bozza del manoscritto. Esse erano inserite manualmente nel sistema di catalogazione locale e i catalogatori dovevano inviare il record CIP agli editori entro due settimane dalla ricezione dei materiali. Dal 1999 gli editori possono compilare direttamente un *form* online allegando la copia digitale delle bozze. Il personale addetto al programma CIP esamina i dati inviati dagli editori al fine di verificarne la completezza e l'integrità. Avviene quindi l'assegnazione del Preassigned Control Number (PCN), ossia un numero di controllo dato alle opere che non sono ancora state pubblicate. La risorsa viene successivamente inviata all'ufficio metadato competente, presso il quale è completata la catalogazione descrittiva fornita dall'editore, vengono assegnate le voci di soggetto basate sulle Library of Congress Subject Headings (LCSH) e i numeri di classificazione relativi alla Library of Congress Classification (LCC) e alla Classificazione decimale Dewey (DDC). Talvolta possono essere aggiunti ulteriori elementi, per esempio, i record creati dal Government Publishing Office (GPO) includono il numero del Superintendent of Documents (SUDOC), mentre quelli creati nell'ambito del Children's and Young Adults' Cataloging Program (CYAC) includono le voci di soggetto CYAC per la narrativa giovanile.

Al termine di questo processo l'editore riceve la descrizione completa che può pubblicare sul verso del frontespizio e, una volta che il libro è stato pubblicato, ne deve essere inviata una copia alla LoC. Il record CIP originale, dopo la pubblicazione del libro, viene arricchito con informazioni quali il numero di pagine e le dimensioni (altezza del dorso in cm). Esso è basato su una versione abbreviata del MARC ed è stato recentemente sviluppato per favorire l'implementazione di RDA, *Resource Description and Access*. Una volta completata la registrazione la LoC distribuisce il record ad altre biblioteche, a distributori e librari, facilitando così il trattamento del libro

5 Agnese Galeffi, *025.3 Catalogazione nella pubblicazione (CIP)*, in *Biblioteconomia guida classificata*, diretta da M. Guerrini, Milano: Editrice Bibliografica, 2007, p. 391.

6 <<https://www.loc.gov/programs/cataloging-in-publication/about-this-program/>>.

con un risparmio notevole di tempo e garantendo la contestuale presenza di registrazioni identiche in tutte le biblioteche.

Nel 2017 il sistema online adottato dalla LoC per le registrazioni CIP è stato completamente trasformato con l'introduzione della piattaforma ServiceNow e dell'applicazione PrePub Book Link (PPBL). Caroline Saccucci, ex responsabile del programma CIP della LoC, ne descrive analiticamente il funzionamento e i vantaggi all'interno dell'articolo *Taking the Library of Congress CIP Program into the Future with PrePub Book Link*⁷, in cui sottolinea come PPBL abbia consentito alla LoC di raggiungere un livello superiore fornendo un servizio nettamente migliorato agli editori e rendendo più efficiente e gestibile il carico di lavoro per i catalogatori interni assegnati al progetto. Grazie alla maggiore interazione dell'applicazione con ONIX editori e catalogatori possono sfruttare appieno i metadati prodotti in questo formato sia per precompilare parte della richiesta CIP sia come parte integrante del flusso di catalogazione. Inoltre, l'editor MARC incluso in PPBL consente la conversione automatica dei dati forniti dall'editore. Nel prossimo futuro è prevista l'integrazione con BIBFRAME.

Il progetto CIP della LoC si concentra sulle monografie statunitensi in corso di pubblicazione, mentre sono esclusi: libri già pubblicati, libri pubblicati al di fuori degli Stati Uniti, libri in lingue diverse da quelle europee occidentali, libri a cui è già stato assegnato un numero di controllo mediante il PCN Program⁸, i libri scolastici al di sotto del livello universitario ad eccezione di quelli relativi alla storia degli Stati Uniti, le edizioni riconfezionate, le traduzioni, i libri elettronici che siano stati pubblicati solo in tale formato, le microforme, i libri tascabili, i periodici, gli spartiti musicali, le guide di viaggio e altre tipologie di risorse.

Possono partecipare solo gli editori che hanno sede legale negli Stati Uniti e che pubblicano titoli suscettibili di esser ampiamente acquistati dalle biblioteche statunitensi. Inoltre, ogni editore deve aver già pubblicato un minimo di tre titoli di tre autori diversi e tutti e tre i titoli devono essere stati acquisiti da un minimo di mille biblioteche statunitensi sia in formato cartaceo sia e-book. Questi dati vengono verificati dai catalogatori coinvolti nel programma tramite WorldCat. Non sono ammessi al progetto i librai, i distributori, le case di produzione o altri

7 Caroline Saccucci, *Taking the Library of Congress CIP Program into the Future with PrePub Book Link*, «JLIS.it», 12 (2021), 3, p. 11-32, <10.4403/jlis.it-12755>.

8 <<https://www.loc.gov/programs/preassigned-control-number/about-this-program/>>.

intermediari nella filiera del libro, gli editori che hanno pubblicato le opere di meno di tre autori e gli editori a pagamento. Essi, tuttavia, non sono esclusi dal progetto gemello PCN che, come anticipato, prevede l'assegnazione di un numero di controllo alle opere non ancora pubblicate e che potrebbero essere aggiunte alle collezioni della LoC, sulla base di una serie di valutazioni condotte dai bibliotecari responsabili dello sviluppo delle raccolte.

Library and Archives Canada

Il programma CIP avviato dalla Library and Archives Canada (LAC)⁹ è messo a disposizione di diverse tipologie di editori canadesi quali: case editrici specializzate, universitarie, di associazioni, di gallerie e musei senza scopo di lucro. Sono, invece, esclusi gli editori che hanno la sede legale in un altro stato, anche nel caso in cui le loro pubblicazioni siano stampate e distribuite in Canada, editori che fanno pagare i servizi redazionali, i self-publishers, i librai, i distributori e qualsiasi altro intermediario. Inoltre, sono escluse dal progetto risorse come i siti web, i blog, i wiki, le pubblicazioni scolastiche (come i manuali), i periodici, le pubblicazioni seriali, le risorse audiovisive, sonore, le versioni digitali di opere pubblicate precedentemente a stampa ecc.

La LAC, analogamente alla Library of Congress, ha elaborato un modello di record, definito CIP data block, da presentare agli editori e costituito da una serie di elementi minimi della descrizione quali: il titolo dell'opera, altri titoli, i nomi (autori persona, enti, curatori, traduttori ecc.), la serie, la descrizione fisica, gli identificatori, i soggetti e la classificazione. Questi elementi rappresentano delle etichette che contrassegnano blocchi di dati sul modello dello schema elaborato dalla LoC nel 2015.

Gli editori coinvolti nel progetto per inviare le proprie richieste alla LAC adottano un *CIP application form*, il quale prevede anche il caricamento delle bozze del frontespizio, della prefazione, dell'introduzione, della copertina e/o dell'indice del libro. Una volta ricevuta la domanda viene presa in carico ed elaborata dall'ufficio CIP. I dati forniti nel modulo di richiesta e nelle prove di stampa allegate vengono utilizzate per creare un record catalografico preliminare successivamente inviato anche all'editore affinché venga inserito nella pubblicazione.

⁹ <<https://library-archives.canada.ca/eng/services/publishers/cataloguing-publication/Pages/cataloguing-publication.aspx>>.

La LAC offre il servizio CIP da oltre quarant'anni e attualmente vi partecipano più di 650 editori. I catalogatori coinvolti creano ogni mese circa 600 nuovi record bibliografici e d'autorità e assicurano la rapida diffusione di questi dati presso le biblioteche canadesi e di tutto il mondo. Le biblioteche possono scaricare i record MARC pagando un abbonamento a OCLC.

British Library

La British Library, dal gennaio 1977, ha avviato un progetto CIP fornendo record bibliografici per opere di prossima pubblicazione edite nel Regno Unito e in Irlanda poi incluse nella British National Bibliography¹⁰. La British Library consente agli editori di inviare i propri record ONIX (ONline Information eXchange) o di compilare un *form* dedicato diviso in cinque sezioni. La prima consente di registrare dati quali: il titolo della risorsa, il sottotitolo, l'edizione, l'ISBN, il numero delle pagine, la lingua della pubblicazione, il numero del volume all'interno della serie (se si tratta di un'opera facente parte di una collana editoriale), il titolo della serie, le dimensioni, le immagini, l'abstract e una serie di informazioni aggiuntive. Di questi dati sono considerati obbligatori solo il titolo dell'opera e l'abstract. Nella seconda sezione è possibile indicare i contributori responsabili della creazione del contenuto artistico e intellettuale dell'opera distinguendo il loro ruolo specifico (autore, curatore, illustratore, traduttore ecc.). Nella terza vengono registrati dettagli relativi al formato e al tipo di supporto della manifestazione, per esempio, viene specificato se è prevista solo una pubblicazione in formato cartaceo o anche elettronico. Nella quarta viene indicato il genere letterario a cui appartiene la risorsa, mentre nella quinta vengono riportati i dati dell'editore.

Il programma CIP consente alla British National Bibliography di fornire informazioni sulle opere di prossima pubblicazione con un anticipo di sedici settimane rispetto alla loro uscita ufficiale. Ciò consente di preannunciare circa 50.000 titoli ogni anno.

Esperienze CIP italiane

In Italia le Biblioteche nazionali centrali non offrono un servizio CIP. La stampa di una scheda catalografica sul verso del frontespizio, che si

10 <<https://www.bl.uk/collection-metadata/metadata-services>>.

trova in alcune pubblicazioni, non corrisponde a un programma CIP nazionale. Gli uffici della Bibliografia nazionale italiana, infatti, a causa dell'allarmante carenza di personale denunciata ormai da alcuni anni, mostrano grandi difficoltà a pubblicare con tempestività i dati relativi alla produzione editoriale nazionale. Per il 2023, per esempio, a oggi sono stati pubblicati solo i primi quattro fascicoli della serie monografie, un fascicolo per la serie periodici, i fascicoli della serie musica a stampa sono fermi al 2022 e quelli relativi alle tesi di dottorato arrivano al primo fascicolo del 2018. Risulta pertanto impossibile svolgere un servizio a supporto degli editori nell'elaborazione delle schede CIP per i volumi in corso di pubblicazione. Al contrario la fornitura di metadati già controllati da parte degli editori alla BNI potrebbe accelerare notevolmente il processo.

In Italia, tuttavia, sono state sviluppate alcune iniziative locali come quella della Biblioteca centrale della Regione siciliana Alberto Bombace e della Biblioteca della Toscana Pietro Leopoldo.

La Biblioteca Alberto Bombace ha avviato una sperimentazione nel 1995, successivamente regolata dalla circolare assessoriale n. 6 del 1999 (Gazzetta Ufficiale della Regione siciliana n. 33 del 1999), per la catalogazione di opere in corso di pubblicazione¹¹. Analogamente a quanto avviene a livello internazionale, la biblioteca predispone un record catalografico che viene poi stampato sui volumi stessi. Sono ammessi al servizio, su base volontaria gli editori siciliani e obbligatoriamente gli organi dall'Assessorato regionale dei beni culturali e dell'identità siciliana e coloro che ricevono finanziamenti da questo ente per iniziative che prevedono la realizzazione di pubblicazioni.

Un servizio analogo è offerto dalla Biblioteca Pietro Leopoldo il cui regolamento del 23 giugno 2020, n. 31 afferma all'articolo 8 comma 6:

La biblioteca collabora all'attività editoriale del Consiglio regionale e della giunta regionale fornendo la sua consulenza e provvedendo alla redazione della Catalogazione nella Pubblicazione (CIP) richiesta dalla struttura competente¹².

La Biblioteca Pietro Leopoldo prosegue in questo modo un'esperienza avviata già alla fine degli anni Settanta del secolo scorso da Luigi Crocetti all'interno della Biblioteca Servizio beni librari della Regione (oggi

11 <<http://www.bibliotecaregionalepalermo.it/index.php?it/282/cip>>.

12 <<https://raccoltanormativa.consiglio.regione.toscana.it/articolo?urndoc=urn:nir:regione.toscana:regolamento.consiglio:2020-06-23;31&pr=idx,0;artic,1;articparziale,0>>.

Biblioteca Crocetti). Crocetti, basandosi sul modello americano, aveva introdotto l'uso di corredare la terza di copertina dei libri che facevano parte delle collane da lui stesso inaugurate (come la collana *Archivi e Biblioteche*) con una tasca all'interno della quale erano inserite tre schede già predisposte per essere inserite nel catalogo per autori, per soggetti e classificato di ogni biblioteca che ne entrasse in possesso. Questa iniziativa, tuttavia, si rivelò fin da subito troppo costosa per poter essere portata avanti.

Analogamente all'Italia altri Stati hanno sviluppato progetti CIP a livello regionale o comunque meno formalizzati. La Biblioteca nazionale di Malta e la Biblioteca dell'Università di Malta, per esempio hanno preso accordi con singoli editori per la fornitura di record CIP, mentre in Spagna questo servizio è offerto da diverse biblioteche locali e universitarie.

Condivisione di dati d'autorità: l'Author Names Project

Accanto ai progetti per la condivisione di record bibliografici alcune realtà hanno avviato una riflessione sulla possibilità di condivisione dei dati d'autorità tra editori e biblioteche. Gli editori, infatti,

intrattengono contatti diretti con i responsabili della creazione delle risorse e sono in grado di fungere da prezioso serbatoio di dati di prima mano e talora esclusivi (per esempio, nuovi autori), quali la data di nascita o il campo di attività di un autore, e rendendo disponibili queste informazioni sul proprio sito web contribuiscono e sostengono il compito identificativo svolto tradizionalmente dalle biblioteche¹³.

Significativa, in questa direzione, è la già citata sperimentazione avviata nell'ambito del Author Names Project, un progetto ambizioso nato con l'obiettivo di scambiare metadati d'autorità con le biblioteche. A tal fine, il Gruppo di lavoro ha sviluppato un'applicazione web open source chiamata Online Author Questionnaire (OAQ) che consente di registrare una serie di informazioni base sull'autore (il nome, gli pseudonimi, altre forme varianti del nome, la data di nascita, l'affiliazione ecc.) attraverso un questionario, secondo una pratica frequente delle case editrici.

I dati inseriti all'interno dell'applicazione OAQ possono essere esportati dagli editori per la creazione di report personalizzati o dalle biblioteche

¹³ Laura Manzoni, *Per una filiera coordinata della metadattazione: gli editori cit.*, p. 93-105.

al fine di mapparli automaticamente e creare dei record d'autorità RDA/NACO, ossia rispondenti ai criteri fissati dalle linee guida RDA e a quelli definiti nell'ambito del Name Authority Cooperative Program.

Tra gli obiettivi del progetto vi era anche la creazione di un protocollo per la trasmissione dei dati al VIAF, *Virtual International Authority File*, tuttavia esso non venne mai realizzato. L'Author Names Project, infatti, è stato avviato nel 2013, ma non è stato successivamente implementato. Esso avrebbe potuto favorire lo sviluppo di iniziative nazionali complementari ai progetti CIP al fine di ottenere descrizioni complete delle risorse e degli autori delle stesse prima della pubblicazione.

Per la condivisione dei dati d'autorità i progetti che oggi svolgono un ruolo centrale sono VIAF e ISNI, *International Standard Name Identifier*, che si inseriscono in più ampi progetti di cooperazione internazionale e authority control. Essi costituiscono

due tra le più importanti iniziative che accettano la sfida di identificare in modo affidabile gli agenti, i nomi geografici ecc. e le opere a essi associate nella rete della conoscenza registrata globale. La loro filosofia si ispira a quella dell'UBC, improntata sull'ideale umanistico della condivisione dei saperi collettivi, della promozione delle diversità culturali e della semplificazione del lavoro delle agenzie bibliografiche e delle biblioteche¹⁴.

Al progetto ISNI, per la creazione di identificatori univoci relativi alle identità pubbliche di persone e organizzazioni partecipano, in qualità di agenzie autorizzate all'assegnazione dei codici identificativi, anche alcuni editori. Per l'Italia, l'agenzia di riferimento, per esempio, è costituita da Casalini Libri, un'azienda di primo piano nella fornitura di servizi bibliografici avanzati e nella distribuzione delle pubblicazioni europee nel mondo¹⁵.

Da ONIX a MARC21: progetti di mappatura e conversione

Accanto allo sviluppo delle forme di cooperazione citate nei paragrafi precedenti, un altro modo per garantire il pieno riutilizzo dei metadati editoriali da parte delle biblioteche è la definizione di modelli e

14 Mauro Guerrini, *Dalla catalogazione alla metadatozione. Tracce di un percorso*, Roma: Associazione italiana biblioteche, 2022, p. 127.

15 Giulia Marangoni, *Casalini Libri diventa Agenzia Italiana per l'ISNI*, «Giornale della libreria», 24 luglio 2017, <<https://www.giornaledellalibreria.it/news-distributore-casalini-libri-diventa-agenzia-italiana-per-lisni-3099.html>>.

pratiche comuni di metadattazione che possano soddisfare entrambe le comunità. Gli elementi base della descrizione realizzata da un editore, infatti, sono gli stessi che compaiono nei record bibliografici. Quindi, per evitare una duplicazione del lavoro le due comunità dovrebbero potersi scambiare reciprocamente dati attraverso l'utilizzo di un linguaggio comune. Per questa ragione sono stati sviluppati progetti di mappatura e conversione dei dati creati a partire dai formati specifici di ogni comunità.

Mentre gran parte delle biblioteche adottano formati quali il MARC21 e l'UNIMARC, le filiere dell'editoria, hanno sviluppato, dalla fine del Novecento, propri formati, tra i quali quello che ha conosciuto maggiore diffusione è ONIX. Si tratta di una famiglia di standard per la descrizione bibliografica e commerciale di libri e riviste, per la registrazione dei diritti di proprietà intellettuale, dei permessi e delle proibizioni d'uso, degli identificatori standard per opere e manifestazioni rendendo le informazioni disponibili in una struttura basata sull'XML che favorisce la condivisione con altre piattaforme. ONIX for Books è stato il primo tra questi standard ed è stato sviluppato nel 2000 in seguito a una riflessione condotta da parte del gruppo EDItEUR insieme al britannico Book Industry Communication e allo statunitense Book Industry Study Group. Attualmente è mantenuto da un International Steering Committee ed è giunto alla versione 3.0, implementata per supportare anche la descrizione degli e-book.

Accanto agli editori, partecipano ai gruppi internazionali di lavoro per lo sviluppo di ONIX anche i rappresentanti del mondo delle biblioteche, i quali hanno colto sin da subito i vantaggi derivanti dalla possibilità di elaborare mappature tra ONIX e formati come il MARC21 o l'UNIMARC. Le agenzie bibliografiche nazionali, responsabili del mantenimento dei programmi CIP, infatti, accettano spesso file ONIX come notifica dei titoli in arrivo, i quali vengono successivamente convertiti in MARC e adeguatamente modificati.

Importanti progetti di mappatura e conversione dei record ONIX in MARC sono stati condotti dalla Library of Congress e da OCLC. Quest'ultimo, in particolare, nel 2009 ha avviato il programma Metadata Service for Publishers con l'obiettivo di produrre record di qualità per editori e biblioteche¹⁶. Attraverso un processo

16 Carol Jean Godby, *Mapping ONIX to MARC*, OCLC, 2010, <<https://www.oclc.org/content/dam/research/publications/library/2010/2010-14.pdf>>; Carol Jean Godby, *A Crosswalk from ONIX Version 3.0 for Books to MARC 21*, OCLC, 2012,

completamente automatizzato sono raccolti i metadati ONIX prodotti dagli editori statunitensi che vengono convertiti in MARC21 e arricchiti dove necessario. Ciò è reso possibile da un software all'interno del quale sono state integrate tabelle di conversione in cui viene specificato che all'elemento <oggetto> di ONIX corrisponde il campo 650 \$a del MARC, all'elemento <autore> corrisponde il campo 100 del MARC ecc.

Il record risultante viene messo a disposizione delle biblioteche e, tramite un flusso separato, agli editori in seguito a una riconversione in ONIX. I dati forniti alle biblioteche devono essere conformi alle AACR2 o alle linee guida RDA, mentre quelli forniti agli editori devono superare il controllo di validazione basato sulle indicazioni del Book Industry Study Group che richiede la presenza di una serie di elementi fondamentali quali: gli identificatori, i titoli, i creatori, gli editori, i soggetti e le dimensioni del supporto. I record finali che vengono presentati a editori e biblioteche, quindi, possono presentare alcune differenze. Ciò dipende dal fatto che nonostante gli elementi base della descrizione adottata da un editore e da una biblioteca coincidano, per gli editori ci sono alcuni dati che rivestono un'importanza maggiore rispetto a quella che gli viene normalmente attribuita da una biblioteca. Si tratta di informazioni che servono a incrementare l'efficienza della catena di fornitura delle pubblicazioni¹⁷:

Il peso di un libro, per esempio, è un'informazione cruciale per massimizzare l'efficienza della logistica. Fondamentali sono anche le informazioni sul prezzo e le condizioni di disponibilità che, tuttavia, sono soggette a continui cambiamenti e richiedono un aggiornamento quotidiano. Questi dati generalmente non compaiono all'interno del catalogo di una biblioteca, che di conseguenza risulta essere molto più statico rispetto a un catalogo editoriale. Un altro aspetto al quale gli editori attribuiscono molta importanza riguarda la registrazione dei metadati semantici che orientano il comportamento d'acquisto dei lettori¹⁸.

<<https://www.oclc.org/content/dam/research/publications/library/2012/2012-04.pdf>>.

17 Piero Attanasio, *New challenges in metadata management between publishers and libraries*, in *Bibliographic control in the digital ecosystem*, a cura di G. Bergamin, M. Guerrini, Firenze: Firenze University Press, 2022, p. 118.

18 Laura Manzoni, *Per una filiera coordinata della metadattazione: gli editori cit.*, p. 99.

Anche relativamente alla classificazione e all'indicizzazione semantica fino alla fine del Novecento, gli editori hanno preso come riferimento schemi comuni al mondo bibliotecario. Tuttavia, a partire dagli anni Duemila hanno sviluppato strumenti di rappresentazione diversi,

più aderenti alle caratteristiche della produzione editoriale presente sul mercato, che fossero capaci al tempo stesso di incontrare le esigenze del commercio librario. In particolare con l'esplosione del commercio online e la moltiplicazione dei canali di accesso ai contenuti offerti dal mondo digitale, nascono standard di classificazione pensati per rappresentare i contenuti librari in modo da tener conto delle modalità di accesso alle informazioni e delle propensioni del pubblico di utenti delle librerie, sia fisiche che virtuali¹⁹.

Il programma di mappatura avviato dalla Library of Congress, a differenza di quello di OCLC è stato presto interrotto rivelandosi fin da subito molto dispendioso a causa dei continui aggiornamenti che avrebbe richiesto. Inoltre, la LoC, come anticipato, si affida al programma CIP per l'acquisizione dei metadati editoriali.

RDA/ONIX Framework for Resource Categorization

Nell'ottobre del 2005 si tenne un incontro tra i membri del Joint Steering Committee for the revision of AACR2²⁰ e i rappresentanti britannici dell'industria editoriale, in occasione del quale entrambe le parti hanno identificato la categorizzazione delle risorse in base al contenuto e al supporto come un interesse reciproco rispetto al quale la cooperazione nella definizione di un modello comune per lo scambio dei dati avrebbe portato notevoli vantaggi. È stata, quindi, avviata un'iniziativa per l'elaborazione di un *framework* capace di favorire lo scambio e il riutilizzo dei metadati prodotti da entrambe le comunità. I lavori sono iniziati nell'ambito di un workshop organizzato a Londra nel marzo del 2006 e hanno portato alla definizione di un modello che identifica due distinti insiemi di attributi che possono essere utilizzati per descrivere una risorsa:

¹⁹ Francesca Cacciapaglia, *Gli standard di Classificazione*, in Anna Lionetti, *Il <dato> è tratto. Guida all'uso dei metadati per l'editoria*, Milano: AIE, 2015, p. 53.

²⁰ Le AACR2 nel 2005 sono state superate e sostituite dalle linee guida RDA.

il primo riguardante il contenuto intellettuale o artistico della risorsa stessa, l'altro il tipo di media e di supporto mediante il quale il contenuto può essere veicolato. Un attributo del contenuto della risorsa, per esempio, è il senso necessario per percepirla (*SensoryMode*) al quale può corrispondere un elenco definito di possibili valori: *vista, udito, tatto, gusto, olfatto* e *nessuno*; mentre un attributo relativo al tipo di supporto è il formato del mezzo di archiviazione (*StorageMediumFormat*) che può essere espresso mediante i valori: *foglio, striscia, rotolo, disco, sfera, cilindro, chip, file server*. L'insieme di questi valori fissi che può essere assunto dall'attributo costituisce un *registro*²¹. I valori utilizzati per descrivere le diverse tipologie di risorse possono essere aggiornati nel corso del tempo all'interno dell'Open Metadata Registry garantendo uno sviluppo delle categorie di RDA coerente con il *framework*.

Il *framework* presenta una grande flessibilità consentendo agli utenti che lo utilizzano di sviluppare dei sottovalori a partire da ciascuno dei valori principali, con la sola limitazione che ogni sottovalore deve appartenere a uno solo dei valori primari. Per esempio, il valore *successione* relativo all'attributo modalità di estensione (*ExtensionMode*) può essere definito attraverso i sottovalori *regolare* e *irregolare*²².

Sebbene la prima versione del *framework* abbia prestato particolare attenzione alla descrizione delle risorse analogiche, è risultata fin da subito evidente la sua applicabilità alle risorse digitali. In particolare, la possibilità di assegnare più categorie e indicare i relativi livelli di applicabilità sarà utile per descrivere oggetti digitali complessi. Inoltre, il *framework* è estendibile ad altre comunità che desiderano migliorare l'interoperabilità dei dati con le biblioteche e gli editori.

Ente NISO per la standardizzazione: raccomandazione di pratiche comuni tra editori e biblioteche

Particolare attenzione è stata rivolta nel corso degli ultimi anni alla condivisione dei metadati relativi alle risorse elettroniche. Per questa ragione l'ente per la standardizzazione NISO ha avvertito la necessità

21 Carlo Bianchini - Mauro Guerrini, *Introduzione a RDA. Linee guida per rappresentare e scoprire le risorse*, Milano: Editrice Bibliografica, 2014, p. 88

22 Gordon Dunsire, *Distinguishing content from carrier. The RDA/ONIX Framework for resource categorization*, «D-Lib Magazine», 13 (January/February 2007), no. 1 /2, <<https://www.dlib.org/dlib/january07/dunsire/01dunsire.html>>.

di elaborare una serie di linee guida, pubblicate nel 2022, all'interno del documento *E-book bibliographic metadata requirements in the sale, publication, discovery, and preservation supply chain*²³, che consentano a editori e bibliotecari di «lavorare su un terreno comune in relazione ai libri digitali, condividendo standard, prassi e soprattutto un linguaggio comune nella predisposizione di dati bibliografici»²⁴. In particolare, le linee guida mirano a soddisfare sei obiettivi fondamentali:

1. definire un set minimo di metadati necessari per descrivere gli e-book al fine di supportare la scoperta, la vendita e la conservazione;
2. individuare il modo più efficace ed efficiente per condividere i metadati attraverso l'intera catena di fornitura;
3. verificare come i record di metadati possano essere aggiornati in seguito al trasferimento delle informazioni;
4. sviluppare un set di regole da seguire per verificare la corrispondenza tra set di metadati che descrivono la stessa risorsa per evitare la duplicazione dei record;
5. fornire esempi di utilizzo di pratiche e standard comuni;
6. condurre delle riflessioni sui formati ONIX e MARC al fine di elaborare eventuali osservazioni.

Il gruppo di lavoro NISO, composto da bibliotecari, editori, aggregatori, distributori ed esperti di conservazione digitale ha cercato di individuare una serie di metadati essenziali che gli editori devono impegnarsi a fornire alle biblioteche nel momento in cui gli vendono degli e-book, ossia:

- titoli
- nomi (di creatori e contributori)
- date
- identificatori dei libri
- soggetti

A essi si possono aggiungere tre ulteriori metadati:

- formato
- limitazioni nell'uso
- URI

²³ NISO, *E-book bibliographic metadata requirements in the sale, publication, discovery, and preservation supply chain. A Recommended Practice of the National Information Standards Organization*. Approvato il 29 gennaio 2022, <<https://groups.niso.org/higherlogic/ws/public/download/26477>>.

²⁴ Rossana Morriello, *Le raccolte bibliotecarie digitali nella società dei dati*, Milano: Editrice Bibliografica, 2020, p. 63.

Nel loro complesso queste informazioni rappresentano un set minimo di metadati essenziali per compiere tre operazioni: trovare e identificare un e-book all'interno di un indice o un elenco; raggruppare i record che si riferiscono allo stesso e-book o a versioni dello stesso e-book; distinguere record che si riferiscono a e-book diversi o a versioni diverse dello stesso e-book.

Conclusioni

L'analisi proposta nell'articolo mostra le diverse modalità di cooperazione che fino a oggi sono state sviluppate per favorire lo scambio di metadati tra editori e biblioteche. Queste comunità pur mostrando interesse reciproco verso i metadati dell'altra continuano a muoversi su binari paralleli facendo riferimento a propri standard e formati di descrizione. Perché possano comunicare è necessario passare attraverso processi di mappatura spesso complessi e dispendiosi anche in ragione dell'evoluzione continua a cui vanno incontro formati ONIX e MARC. Per superare in modo definitivo questo ostacolo è necessario giungere alla definizione di un set di metadati condiviso, capace di rispondere alle esigenze di entrambe le comunità, ed elaborare un modello di dati unico indipendente da uno specifico dominio come RDF (Resource Description Framework) utilizzato per la strutturazione dei linked data. Un contributo decisivo in questa direzione può venire da RDA, il primo standard internazionale rivolto a chiunque abbia la necessità di descrivere una risorsa bibliografica compresi gli editori. Esso, al contrario delle AACR2 che erano state concepite per operare con il formato MARC, presenta una flessibilità e una granularità che fanno dei linked data la tecnologia più adatta alla sua applicazione.

Bibliografia

- Piero Attanasio, *New challenges in metadata management between publishers and libraries*, in *Bibliographic control in the digital ecosystem*, a cura di G. Bergamin, M. Guerrini, Firenze: Firenze University Press, 2022, p. 116-122.
- Carlo Bianchini - Mauro Guerrini, *Introduzione a RDA. Linee guida per rappresentare e scoprire le risorse*, Milano: Editrice Bibliografica, 2014.
- Gordon Dunsire, *Distinguishing content from carrier. The RDA/ONIX Framework for resource categorization*, «D-Lib Magazine», 13 (January/

- February 2007), no. 1 /2, <<https://www.dlib.org/dlib/january07/dunsire/01dunsire.html>>.
- Agnese Galeffi, *025.3 Catalogazione nella pubblicazione (CIP)*, in *Biblioteconomia guida classificata*, diretta da M. Guerrini, Milano: Editrice Bibliografica, 2007, p. 391.
- Carol Jean Godby, *A Crosswalk from ONIX Versione 3.0 for Books to MARC 21*, OCLC, 2012, <https://www.oclc.org/content/dam/research/publications/library/2012/2012-04.pdf>.
- Carol Jean Godby, *Mapping ONIX to MARC*, OCLC, 2010, <<https://www.oclc.org/content/dam/research/publications/library/2010/2010-14.pdf>>.
- Fulvio Guatelli, *FUP Scientific Cloud e l'editoria fatta dagli studiosi*, «Società e storia», 167 (2020), p. 155-164.
- Mauro Guerrini, *Dalla catalogazione alla metadattazione. Tracce di un percorso*, Roma: Associazione italiana biblioteche, 2022.
- IFLA, *Common Practices for National Bibliographies in the Digital Age*. 14 luglio 2022, <<https://repository.ifla.org/bitstream/123456789/2001/1/Bibliographies%20Common%20Practices%202022.pdf>>.
- IFLA, *Pratiche condivise per le bibliografie nazionali nell'era digitale*, edizione italiana a cura dell'AIB, Commissione nazionale biblioteche e servizi nazionali. 2023, <<https://repository.ifla.org/bitstream/123456789/2582/1/common-practices-for-national-bibliographies-in-the-digital-age-it.pdf>>.
- Deirdre Kiorgaard, *RDA/ONIX Framework for Resource Categorization*. 3 agosto 2006, <<https://www.loc.gov/marc/marbi/2007/5chair10.pdf>>.
- Anna Lionetti, *Il <dato> è tratto. Guida all'uso dei metadati per l'editoria*, Milano: AIE, 2015.
- Laura Manzoni, *Per una filiera coordinata della metadattazione: gli editori*, in Mauro Guerrini, *Metadattazione*, Milano: Editrice Bibliografica, 2022, p. 93-105.
- Giulia Marangoni, *Casalini Libri diventa Agenzia Italiana per l'ISNI*, «Giornale della libreria», 24 luglio 2017, <<https://www.giornaledellalibreria.it/news-distributori-casalini-libri-diventa-agenzia-italiana-per-lisni-3099.html>>.
- Rossana Morriello, *Le raccolte bibliotecarie digitali nella società dei dati*, Milano: Editrice Bibliografica, 2020.
- Network Development and MARC Standards Office Library of Congress, *ONIX to MARC 21 Mapping*, <<https://www.loc.gov/marc/onix2marc.html>>.
- NISO, *E-book bibliographic metadata requirements in the sale, publication, discovery, and preservation supply chain. A Recommended Practice of the National Information Standards Organization*. Approvato il 29 gennaio 2022, <<https://groups.niso.org/higherlogic/ws/public/download/26477>>.
- Caroline Saccucci, *Taking the Library of Congress CIP Program into the Future with PrePub Book Link*, «JLIS.it», 12 (2021), no. 3, p. 11-32, <10.4403/jlis.it-12755>.

