

*Le ontologie bibliografiche. Modelli concettuali e vocabolari condivisi per l'universo bibliografico*, a cura di Maria Teresa Biagetti, Roma, Bulzoni, 2022, 161 p., ill., (Biblioteconomia Scienza dell'informazione Archivistica, 1), ISBN 978-88-6897-280-6, € 18,00.

Il volume *Le ontologie bibliografiche. Modelli concettuali e vocabolari condivisi per l'universo bibliografico* a cura di Maria Teresa Biagetti è il primo della collana Biblioteconomia, Scienza dell'informazione, Archivistica diretto dalla stessa Biagetti per l'editore Bulzoni. Questa elegante pubblicazione raccoglie sei contributi di tre autrici dedicate al tema dei modelli concettuali e delle ontologie prodottisi negli ultimi anni e di interesse per l'universo bibliografico e, in particolare, lo mettono in relazione anche con lo sviluppo del web semantico.

Anche se i soggetti dei contributi che confluiscono nella raccolta sono diversi, tutti sono collegati mediante una classe generale rappresentata proprio dalle ontologie bibliografiche – da cui il titolo complessivo del volume – che rappresentano concetti e classi dell'universo del discorso e le relazioni di base che tra di essi intercorrono.

L'obiettivo del libro è quindi «rendere conto delle diverse tipologie di ontologie che riguardano l'universo bibliografico, presentare e discutere le caratteristiche dei più noti e fondamentali modelli concettuali, partendo da una riflessione fondamentale sulle relazioni bibliografiche, che sono state per secoli il cuore della trattazione catalografica» (p. 10-11).

È perciò evidente la ragione per cui proprio il saggio di apertura

del volume sia dedicato da Maria Teresa Biagetti a *Le relazioni bibliografiche*. Maria Teresa Biagetti insegna Biblioteconomia e Scienza dell'informazione all'Università La Sapienza di Roma e, per tutta la sua carriera, si è occupata di indagare gli aspetti teoretici della catalogazione bibliografica, dell'indicizzazione semantica e dell'organizzazione della conoscenza registrata (Knowledge Organization), con uno specifico interesse verso l'evoluzione di questi campi di studio e l'impatto che su di essi ha avuto l'avvento delle tecnologie informatiche. In particolare, l'interesse dell'autrice verso gli aspetti tecnologici è ben testimoniato da due recenti monografie dedicate a *La Scienza dell'informazione* e a *Le biblioteche digitali*.

Le relazioni bibliografiche rappresentano un tema e un problema centrale nella teoria della catalogazione e la definizione delle relazioni è un argomento dibattuto da secoli, che affonda le sue radici nei principi enunciati due secoli fa da Giovanni Battista Audiffredi, che aveva sottolineato l'importanza dell'organizzazione bibliografica delle opere e delle loro edizioni. Già con Audiffredi il concetto di opera si distingue da quello di edizione e si delinea come creazione intellettuale autonoma comprendendo così testi di qualsiasi estensione come le lettere e le parti di opere. Biagetti ricostruisce con puntualità la lunga tradizione della riflessione sulle opere, attraverso il pensiero di Panizzi, evidente nell'indicizzazione di singoli contributi all'interno del catalogo classificato delle opere scientifiche possedute dalla biblioteca della Royal Society, realizzato nel 1936. Biagetti identifica poi nei *Cataloguing rules and principles* di Lubetzky il capostipite della teorizzazione dell'individuazione di tutte le edizioni e le traduzioni di un'opera nella costruzione dei cataloghi bibliografici. La ricostruzione storica, attenta ed accurata, si sofferma anche sui contributi fondamentali dei *Functions and objects of author and title cataloguing* di Domanovsky (disponibile anche nell'edizione italiana), di *The catalogue as access mechanism* di Wilson e delle *Bibliographic relationships* di O'Neill e Vizine-Goetz. Un contributo decisivo nel campo delle relazioni bibliografiche è stato dato però da Barbara B. Tillett in una im-

portantissima tesi di dottorato che, partendo dalla rappresentazione bidimensionale e ‘piatta’ del formato MARC allora disponibile, dimostra la necessità di sviluppare una tassonomia più estesa e dettagliata delle relazioni tra le unità bibliografiche. Il lavoro di Tillett è stato la base degli sviluppi successivi, sia con i contributi di Smiraglia e Leazer che con quello decisivo di FRBR, che ha orientato e guidato tutta la riflessione seguente, fino a RDA (*Resource Description and Access*) e IFLA Library Reference Model (LRM). La rassegna individua come punto di conclusione ideale, necessariamente provvisorio, di questa evoluzione nel CIDOC-CRM e in FRBRoo, che adottano un approccio alle relazioni bibliografiche orientato non alle entità, ma agli eventi che portano alla loro esistenza (come un evento di produzione di un supporto, o di ideazione di un’opera, o di creazione di un’espressione e così via).

Il secondo saggio, sempre a firma di Maria Teresa Biagetti, introduce all’altro tema fondamentale del volume che stiamo presentando: quello delle ontologie. Biagetti – che è membro storico del capitolo italiano dell’ISKO – sottolinea che «nel settore della scienza dell’informazione, e in particolare nel settore specifico denominato Library and Information Science (LIS), il focus non è sull’applicazione del ragionamento logico, ma sui documenti che descrivono le entità del mondo reale e immaginario dei diversi domini e secondo molteplici punti di osservazione» (p. 26). Biagetti si interroga sul rapporto tra ontologie nell’ambito della Computer Science e le tassonomie biblioteconomiche, perché alcuni – come Ingetraut Dahlberg – ritengono che il termine ontologie sia soltanto una nuova denominazione per quelle realizzazioni che nel settore della LIS sono note come classificazioni bibliografiche e biblioteconomiche. La questione rimanda alla definizione di ontologia, della quale Biagetti offre un approfondito excursus che evidenzia la duplice necessità da un lato, della formalizzazione, e dall’altro quella dell’individuazione precisa delle relazioni che culmina con una proposta di definizione che possa rendere conto di questi aspetti: «ontologies are a kind of KOS that present the hi-

ghest degree of semantic richness, as they allow to establish a great number of relations between terms, and provide attributes for each class» (p. 33).

Biagetti sottolinea che lo scopo fondamentale delle ontologie è consentire l'interoperabilità di tipo semantico tra sistemi, secondo la definizione proposta da Marcia Zeng, per cui «l'interoperabilità semantica può essere definita come la capacità che hanno diversi agenti, servizi e applicazioni *to communicate (in the form of transfer, exchange, transformation, mediation, migration, integration, etc.) data, information, and knowledge – while ensuring accuracy and preserving the meaning of that same data, information, and knowledge*» (p. 34).

Il terzo saggio di questo volume a tre nomi è dedicato da Irene Piccari a *BibFrame e le relazioni bibliografiche nel web dei dati* (p. 41-88). Irene Piccari ha approfondito lo studio delle ontologie nell'ambito della Scienza dell'informazione durante il suo percorso di studi magistrale in Archivistica e biblioteconomia, prestando particolare attenzione proprio a quelle relative al dominio bibliografico. Dopo essersi specializzata in Gestione e conservazione di archivi digitali (Università di Macerata), ha lavorato in diverse aziende informatiche in progetti di digitalizzazione dei processi e degli asset informativi e oggi si occupa della definizione di ontologie per la modellazione di specifici ambiti informativi e dell'applicazione delle tecnologie del web semantico allo sviluppo di progetti linked open data.

In questo contributo l'autrice, dopo una «trattazione di respiro più ampio volta a presentare il modello concettuale» (p. 41) e la «disamina delle caratteristiche dell'ontologia e delle sue evoluzioni intercorse nelle diverse fasi del progetto» (p. 41), si dedica a un esame dettagliato della totalità delle classi e delle relazioni modellate da BibFrame, allo scopo principale di analizzare le possibilità espressive offerte da BibFrame stesso per la rappresentazione delle relazioni bibliografiche. Il contributo è corredato da un ricco apparato iconografico, al termine del quale hanno un ruolo importante due tabelle (n. 8 e n. 9) dedicate rispettivamente alle *Classi definite nel BibFrame Vocabulary*

*lary* (p. 70-75) e alle *Proprietà definite nel BibFrame Vocabulary* (p. 76-83). La prima mira a evidenziare e agevolare il reperimento delle informazioni relative alle relazioni gerarchiche tra le classi e al loro numero; la seconda mostra le medesime informazioni relativamente alle proprietà definite nel BibFrame Vocabulary. In entrambe le tabelle il rapporto di subordinazione gerarchica tra gli elementi in relazione è espresso mediante la posizione di un elemento nella seconda, terza o quarta colonna, che indicano rispettivamente il Livello 1, il Livello 2 o il Livello 3.

È firmato da Irene Piccari anche il quarto contributo dal titolo *IFLA Library Reference Model. Il modello concettuale e un confronto con BibFrame* (p. 89-110). La scelta di descrivere IFLA LRM nasce dall'interesse dell'autrice per «la vocazione del nuovo modello ad essere utilizzato come base per l'interpretazione delle entità bibliografiche nell'ambiente aperto del web» (p. 90). Il testo quindi procede seguendo un percorso logico che dapprima prevede l'analisi del modello IFLA LRM, con un breve excursus sui principali testi italiani di riferimento – ai quali si può aggiungere, a vantaggio del lettore, anche la segnalazione della traduzione italiana di IFLA LRM curata dall'ICCU, che l'autrice riporta soltanto nei riferimenti bibliografici finali – e poi si concentra sulle relazioni definite tra le entità del Gruppo 1 di FRBR (ovvero Opera, Espressione, Manifestazione e Item), ma senza entrare nel merito delle differenze o della continuità di queste relazioni rispetto ai modelli di origine (ovvero FRBR, FRAD e FR-SAR). Nella seconda parte del saggio, l'autrice procede al confronto puntuale tra l'espressività delle relazioni in IFLA LRM e in BibFrame, senza entrare nel dettaglio della «analisi di tutte le entità e le relazioni modellate» (p. 90) ma focalizzandosi su quelle che il modello stesso definisce *core*, ovvero essenziali per soddisfare i requisiti minimi del modello stesso. Il confronto tra i due modelli viene svolto esaminando tre diversi approcci: i *livelli di espressività* utilizzati nella costruzione dei modelli e gli strumenti informativi utili, la *struttura delle classi* principali che derivano dalle scelte di modellizzazione e di definizione

delle risorse bibliografiche e, infine, le *proprietà*, analizzate nella loro costruzione, nella definizione e nei vincoli a cui sono sottoposti. Con l'aiuto di alcune tabelle, l'autrice riesce efficacemente a «evidenziare le possibilità espressive previste nei due modelli, utili al collegamento tra istanze appartenenti alla stessa classe che intrattengano tra loro relazioni significative per la descrizione o contestualizzazione delle risorse bibliografiche» (p. 98). Il confronto è sintetizzato nelle tabelle 2 e 3.

La tabella 2 (p. 96) evidenzia che l'apparato informativo di IFLA LRM è più completo, perché include anche indicazioni sulla cardinalità, le note d'ambito e gli esempi.

La tabella 3 (p. 101-103) presenta la mappatura tra le classi principali di BibFrame e IFLA LRM e consente con un solo colpo d'occhio di cogliere l'assenza in BibFrame di tre entità: RES, sovraordinata gerarchicamente a tutte le altre, ESPRESSIONE e NOMEN. Quest'ultima sembra tra tutte la differenza più sostanziale, per l'importanza che la gestione dei NOMEN ha non solo nella catalogazione classica ma soprattutto nella gestione delle identità nel web semantico.

Ne emerge inoltre che il gruppo di relazioni ricorsive per la classe Opera risulta più numeroso nel *BibFrame Vocabulary*, ma anche che IFLA LRM definisce in dettaglio quelle relazioni eludendo possibili margini di ambiguità semantica, perché specifica sempre il dominio e il codominio delle relazioni. Irene Piccari conclude che i due modelli mostrano una convergenza nella definizione delle relazioni tra risorse ed evidenzia come essi siano stati progettati secondo una prospettiva che in futuro diventerà la norma, ovvero verso le «concrete implementazioni di utilizzo» anziché la sola progettazione di modelli bibliografici integrabili nel web dei dati.

*Le ontologie bibliografiche nei progetti Linked Open Data delle biblioteche europee* è il titolo del quinto saggio del volume, firmato da Antonella Iacono (p. 111-148). Diplomata alla Scuola speciale per Archivisti e Bibliotecari (SSAB) di Roma, dottore di ricerca in Scienze librerie e documentarie presso Università di Roma La Sapienza

(2013), membro del Gruppo di studio nazionale dell'AIB "Catalogazione, open data e web semantico CILW" per il triennio 2014-2017, Antonella Iacono ha mostrato un ampio raggio di interessi scientifici: dallo sviluppo degli OPAC bibliotecari e delle biblioteche digitali, alla *Knowledge Organization*, dalle teorie del comportamento informativo (*information behavior e information seeking behavior*) all'*information retrieval*. In particolare, ha sviluppato conoscenze e competenze nell'ambito del Web semantico, degli open data e dei linked data in campo bibliotecario, sui quali ha pubblicato alcuni saggi e ha tenuto corsi di docenza integrativa presso la Scuola di specializzazione in Beni archivistici e librari di Sapienza Università di Roma in questo settore di studi.

Il focus del contributo di Antonella Iacono è sui progetti LOD di grandi dimensioni che siano stati sviluppati da rilevanti istituti bibliotecari europei. Infatti, le potenzialità offerte dal web semantico per il mondo bibliotecario sono state evidenziate sin dall'esordio dei Linked data e si possono convenzionalmente fare risalire al primo *Report* pubblicato dal Library Linked Data Incubator Group nel 2011, che già allora segnalava l'altissimo valore dei dati di autorità delle biblioteche per lo sviluppo del web semantico e indicava i primi progetti avviati in ambito bibliotecario. Non c'è dubbio che ogni progetto di successo diventa inevitabilmente «fonte di ispirazione per altre iniziative» (p. 111), soprattutto quando è realizzato in ambito europeo e con proposte di rilievo. Nel 2018, ricorda Antonella Iacono, erano «oltre 140 le istituzioni nel mondo attivamente impegnate in progetti LOD, di cui una considerevole parte in Europa [...] ed erano] molte le istituzioni che hanno guardato ai LOD come esito possibile del catalogo e degli strumenti di ricerca bibliografica, in un decennio peraltro ricco di novità sul fronte dei linguaggi, standard e modelli concettuali che sono stati riformulati e ripensati per il Web semantico» (p. 111).

Da questo punto di vista, è condivisibile la scelta dell'autrice di un approccio che studi l'attività di modellazione semantica in relazione al contesto di sviluppo, di risultati e di potenzialità applicative che

emergono da «cinque progetti significativi di British Library, Kungliga biblioteket, Bibliothèque Nationale de France, Biblioteca Nacional de Espana e Europeana Foundation». Infatti, scrive Iacono che «al centro del flusso di lavoro di produzione dei LOD, la modellazione semantica comporta la scelta delle ontologie e degli schemi più appropriati a descrivere nel Web il proprio set di dati» (p. 112). Si tratta di un'operazione tutt'altro che meccanica; anzi, come sottolinea l'autrice, si tratta di «un processo estremamente creativo nel quale si dà forma a nuove modalità interpretative del processo di ricerca informativa, che consentono [...] una grande sperimentazione sulle interfacce di ricerca allo scopo di creare nuova conoscenza» (p. 112).

Il processo è creativo anche perché le sole ontologie di dominio non sono sufficienti a esaurire le esigenze della modellazione, perché si dovrà prevedere «il riutilizzo di vocabolari e ontologie esistenti per favorire la massima interoperabilità e riutilizzo dei dati» (p. 113). Perciò le diverse iniziative mostrano che «a seconda degli obiettivi dell'iniziativa, della portata e dell'uso previsto delle risorse» ciascuna «ha sviluppato la propria interpretazione di come le proprie risorse possono o debbano essere integrate nel Web semantico» (p. 116).

È chiaro a questo punto il valore dell'analisi del processo di modellazione dei dati nel quadro dei cinque casi europei ricordati poco sopra che si può trovare all'interno del contributo. Sarebbe troppo lungo riproporre qui in breve anche uno soltanto dei casi scelti, ma vale la pena senz'altro di ricordare alcune conclusioni che l'autrice trae dall'analisi.

Iacono scrive che «conseguentemente allo sviluppo in ambito bibliografico di RDA, un numero [sic] di progetti utilizza già FRBR come modello concettuale di riferimento e come modello per la struttura dei dati» (p. 142). Anche se i progetti presi in esame hanno avuto tempi di sviluppo e, inevitabilmente, esiti diversi, Iacono sottolinea che «in ogni progetto citato la fase della modellazione semantica ha comportato ben precise scelte che si riflettono sull'interoperabilità, sulle possibilità di scoperta e sulle funzionalità possibili nelle interfac-

ce di ricerca» (p. 143). Questa interoperabilità deve poi essere intesa «non solo in senso tecnologico ma anche semantico, in quanto [...] in grado di potenziare le zone di sovrapposizione di una enorme ricchezza informativa e di permettere il riutilizzo di queste risorse digitali in ambiti diversi» (p. 143).

Si aprono scenari molto complessi, e non si può che condividere l'opinione dell'autrice quando afferma che «è ancora prematuro intravedere gli sviluppi futuri delle interfacce, le quali dovranno sempre di più prevedere un utilizzo più ampio possibile dei dati prodotti» (p. 145).

L'ultimo contributo, a firma della curatrice del volume, è dedicato a *Vocabolari condivisi per le entità bibliografiche: Bibo e le ontologie del gruppo SPAR* (p. 149-161). Un articolo sulle ontologie a fine volume è utile perché le «ontologie ambiscono a definire la conoscenza espressa in un dominio scientifico particolare al fine di agevolarne la comprensione» (p. 149).

Avere a disposizione le ontologie per le entità bibliografiche offre quindi il vantaggio di fornire una mappa del dominio, che è di fondamentale importanza per comprendere, gestire e produrre dati in quel particolare dominio.

La prima ontologia presa in esame è Bibo, sviluppata nel 2009 per «definire le entità bibliografiche nel semantic web» ed è in uso da parte della Deutsche Nationalbibliothek e della British National Bibliography, in entrambi i casi insieme a Dublin Core. Biagetti fa notare che una delle caratteristiche apprezzabili dell'ontologia Bibo è «la cura che gli sviluppatori hanno riposto nella definizione delle classi e delle sottoclassi che appartengono al settore bibliografico (p. 149) alla quale però fa da rovescio della medaglia il numero esiguo di proprietà applicabili al settore bibliografico presenti in questa ontologia» (p. 149-150).

Oltre a Bibo, l'autrice prende in considerazione un gruppo di ontologie complementari, sviluppate congiuntamente dall'Università di Bologna e quella di Oxford, che si chiamano SPAR, ovvero *Semantic*

*Publishing and Referencing Ontologies*, e che mirano a modellare semanticamente i dati relativi alle pubblicazioni scientifiche nei diversi campi di studio, per facilitare la circolazione delle informazioni relative alla letteratura scientifica. In particolare, le ontologie SPAR si dedicano alla definizione di un modello per il *semantic publishing* finalizzato a registrare dati semantici di un articolo relativi alla sua pubblicazione, o utili per il suo ritrovamento, per il reperimento dei dati contenuti in esso, per il collegamento dell'articolo ad altri simili e, più in generale, per la comprensione del suo contenuto. La più nota delle 12 ontologie SPAR è FaBiO, ovvero *FRBR-aligned Bibliographic Ontology*. Questa contempla entità pubblicate o potenzialmente pubblicabili come articoli, monografie e pubblicazioni in serie, ma anche pagine web, set di dati, cataloghi e algoritmi. Biagetti mostra, ricorrendo a una tabella (1. *Mappatura tra alcune classi in FaBiO e categorie bibliografiche in FRBR e in IFLA LRM*) che l'allineamento tra FaBiO, FRBR e IFLA LRM non è perfetto e fa notare che dal confronto emerge in modo evidente che FaBiO non chiarisce del tutto la differenza tra Opera e Espressione. Nell'analisi successiva (tabella 2), l'autrice rileva che né FaBiO né Bibo sono in grado di rappresentare risorse derivate da altre risorse. Tuttavia, l'autrice segnala che a suo avviso «la più significativa ontologia del gruppo SPAR è CiTO, Citation Typing Ontology», che definisce i «metadati che possono essere utilizzati per descrivere le citazioni bibliografiche» mettendo in evidenza le intenzioni e le «motivazioni che sono alla base delle citazioni stesse» (p. 156).

Credo che – al di là degli aspetti tecnici sui quali l'articolo entra davvero molto in dettaglio – la mancata modellazione delle relazioni derivate da un lato e l'importanza giustamente attribuita a CiTO siano indicazioni che portano a evidenziare l'attenzione delle ontologie del gruppo SPAR verso la mappatura della letteratura scientifica corrente, più che retrospettiva, e la rilevanza che in questa mappatura assumono i fenomeni della *citation analysis*, della *co-citation*, della *distant citation* e della *self citation*.

Per concludere, si può senz'altro affermare che il volume raggiunge

pienamente l'obiettivo di presentare le numerose e diverse ontologie di interesse per l'ambito bibliografico e inoltre ne offre un'analisi attenta e critica. Oltre a un approccio teorico solido, che riallaccia con chiarezza i modelli più recenti a una tradizione di studi catalografici di lunga durata, il volume procede dal punto di vista metodologico con dettagliate comparazioni tra le diverse ontologie, consentendo al lettore di cogliere sempre con rapidità i punti di forza e i punti di debolezza di ciascuna. Infine, la segnalazione di casi applicativi permette di orientarsi ancora meglio tra le numerose opzioni scelte e fornisce la base per l'avvio di mappature che richiedono – come il volume dimostra chiaramente – di contemperare inevitabilmente conoscenze teoriche, conoscenze tecniche e creatività.

*Carlo Bianchini*