

Identifiers within conceptual models and standards

Laura Manzoni^(a)

a) University of Florence, <https://orcid.org/0000-0002-2883-7916>

Contact: Laura Manzoni, lauramanzoni1993@live.it

Received: 26 April 2022; **Accepted:** 16 June 2022; **First Published:** 15 September 2022

ABSTRACT

Identifiers are numeric or alphanumeric strings of characters permanently and uniquely associated with a resource to ensure stable access to it and its metadata. They play a fundamental role in the semantic web as they guarantee interoperability between different systems. Since their emergence, libraries and other institutions have been aware of their importance in data quality control processes, for the realization of authority control and universal bibliographic control, and have dedicated specific chapters for their registration within conceptual models, cataloguing standards and have provided specific fields in cataloguing formats.

KEYWORDS

Identifiers; Standards; Conceptual Models; Cataloguing Standards.

Gli identificatori all'interno di modelli concettuali e standard

ABSTRACT

Gli identificatori sono stringhe numeriche o alfanumeriche di caratteri associate in modo permanente e univoco a una risorsa per garantire l'accesso stabile a essa e ai suoi metadati. Svolgono un ruolo fondamentale nel web semantico poiché garantiscono l'interoperabilità tra sistemi diversi. Fin dalla loro comparsa le biblioteche e le altre istituzioni della memoria registrata sono state consapevoli della loro importanza nei processi di controllo di qualità dei dati, per la realizzazione dell'authority control e del controllo bibliografico universale e hanno dedicato a essi capitoli specifici per la loro registrazione all'interno di modelli concettuali, standard catalogafici e hanno previsto campi appositi nei formati di catalogazione.

PAROLE CHIAVE

Identificatori; Standard; Modelli concettuali; Standard di catalogazione.

Gli identificatori

Gli identificatori sono delle stringhe numeriche o alfanumeriche di caratteri associati in maniera univoca alle risorse per garantire l'accesso stabile a esse. Sono normalmente attribuiti e registrati da agenzie bibliografiche nazionali e da istituzioni commerciali. Essi devono essere costruiti sulla base di alcuni principi quali l'unicità e la permanenza affinché possano svolgere le loro funzioni in modo efficace. L'unicità determina che a ogni identificatore si riferisca una sola risorsa mentre la permanenza garantisce la stabilità di un identificatore nel tempo. Fondamentale è, inoltre, l'univocità che rende un'entità inequivocabilmente riconoscibile all'interno di uno specifico contesto di produzione, di un dominio.

Gli identificatori hanno sempre svolto un ruolo fondamentale in ambito informatico in tutti i sistemi di elaborazione delle informazioni e, in particolare, nelle basi dati relazionali. A partire dagli anni Settanta del Novecento hanno cominciato a essere utilizzati dagli editori per identificare in maniera univoca le proprie pubblicazioni o parti di esse. Nel 1967 venne introdotto in Inghilterra e negli Stati Uniti SBN (Standard Book Number) da cui derivò l'ISBN (International Standard Book Number) oggi utilizzato per l'identificazione dei libri e di altre risorse come le carte geografiche, i cd, i dvd, gli audiolibri ecc. Nel 1975 nacque l'ISSN (International Standard Serial Number) per le pubblicazioni seriali. Successivamente vennero introdotti identificatori rivolti ad altre tipologie di prodotti editoriali come l'ISRC (International Standard Recording Code) e l'ISMN (International Standard Music Number) che identificano oggetti definibili su base materiale e consolidati sul mercato. Già nel corso degli anni Novanta la comunità dell'informazione decise che era arrivato il momento di identificare anche le opere intellettuali indipendentemente dalla loro forma di presentazione. Nacquero così identificatori come l'ISWC (International Standard Musical Work Code) e l'ISTC (International Standard Text Code). Tuttavia, fu soprattutto in seguito allo sviluppo del digitale che si avvertì l'urgenza di identificare nuove tipologie di risorsa, i loro metadati e i loro link (Vitiello 2004, p. 70). Vennero quindi elaborati identificatori come il DOI (Digital Object Identifier), l'NBN (National Bibliography Number) e l'ARK (Archival Resource Key). Accanto a essi, infine, sono stati sviluppati progetti specifici volti all'identificazione di persone, famiglie ed enti pubblicamente conosciuti nella filiera della produzione intellettuale, culturale e scientifica. Tra i più famosi vanno ricordati il VIAF (Virtual International Authority File), l'ISNI (International Standard Name identifier) e ORCID (Open Research Contributor ID) (Concedda 2015).

L'utilizzo degli identificatori in ambito catalografico

Nel corso dell'ultimo decennio gli identificatori creati dalle diverse agenzie bibliografiche si sono moltiplicati e hanno assunto un ruolo fondamentale nella condivisione dei dati bibliografici. L'identificazione, infatti, costituisce l'attività caratterizzante il web semantico all'interno del quale ogni oggetto è rappresentato mediante un URI (Uniform Resource Identifier) che rende la risorsa disponibile come link utilizzabile e dereferenziabile. La presenza delle biblioteche nel web e lo sviluppo incontrollato dell'editoria digitale, quindi, hanno reso gli identificatori strumenti fondamentali nei processi di controllo della qualità dei dati, per la realizzazione dell'authority control e del controllo bibliografico universale. Il controllo d'autorità è sempre stato re-

alizzato attraverso la scelta di una forma autorizzata del nome dell'entità descritta (una persona, un ente, un'opera ecc.), ma da quando le procedure di catalogazione sono state automatizzate, a partire dagli anni Sessanta, è apparso chiaro il vantaggio di utilizzare numeri di controllo anziché stringhe testuali vincolate a convenzioni linguistiche e culturali. Progetti internazionali come VIAF e ISNI sono una chiara testimonianza di questa trasformazione (Moi 2021, p. 376). Essi consentano l'aggregazione di dati provenienti da fonti diverse in cui ogni etichetta rappresenta una forma del nome associata a un'entità in una determinata lingua e scrittura ed è valida in uno specifico contesto di applicazione. Le etichette vengono riunite mediante l'uso di identificatori numerici univoci utilizzati dalle macchine per arrivare alla descrizione della risorsa e consentendo al lettore di scegliere una delle forme varianti del nome dell'entità comprese nel medesimo cluster. Se, per esempio, facciamo una ricerca sul VIAF con il nome Giacomo Leopardi viene restituito il seguente cluster (un insieme, un aggregato) di forme varianti del nome dell'entità utilizzate da diverse agenzie bibliografiche seguite dal VIAF ID, ossia l'identificatore univoco associato a questa entità:

- Leopardi, Giacomo, 1798-1837
- Leopardi, Giacomo
- Leopardi, Giacomo, conte, 1798-1837
- Γζάκομο Λεοπάρντι
- יִדְרֵפּוּאֵל, וּמוֹקֵג, 1798-1837
- Леопарди, Дж. 1798-1837 Джакомо
- Leopardi, G. 1798-1837 Giacomo
- VIAF ID: 12311353 (Personal)

Per sfruttare al meglio le potenzialità degli identificatori le agenzie bibliografiche nazionali stanno impegnandosi nel definire delle *policy* specifiche per la loro gestione, tenendo conto del quadro generale delle attività da compiere sia all'interno della biblioteca, integrando la gestione degli identificatori nel lavoro svolto quotidianamente dai catalogatori, sia all'esterno in relazione all'uso degli identificatori da parte degli utenti finali. In particolare, è necessario tener conto della varietà degli identificatori che ogni biblioteca deve gestire e del tipo di responsabilità che può esercitare su di essi: se può occuparsi direttamente della loro attribuzione o se deve limitarsi alla loro registrazione inviando i dati necessari per chiederne l'attribuzione all'agenzia internazionale che li gestisce (Boulet 2022).

Gli identificatori all'interno dei modelli concettuali e negli standard di catalogazione

Gli identificatori, a partire dagli anni Settanta del secolo scorso, hanno acquisito un'importanza sempre maggiore in ambito catalografico. Nella prima edizione di ISBD(M), *International Standard Bibliographic Description for Monographic Publications* del 1974 era prevista un'area specifica, l'area 7, dedicata alla registrazione dell'ISBN. Nel 1977 con la pubblicazione di ISBD(S), *International Standard Bibliographic Description for Serials* venne introdotta una nuova area dedicata alle diverse tipologie di risorsa, che divenne l'area 3, mentre la precedente area 3 divenne l'area 4 e tutte le altre

scalarono di una cifra modificando la struttura dello standard. L'identificatore, che in questo caso era l'ISSN, veniva trattato in area 8. Sempre nel 1997 uscì ISBD(G), *International Standard Bibliographic Description General* con l'obiettivo di costituire una guida generale a cui avrebbero dovuto uniformarsi tutti gli standard successivi relativi alle diverse tipologie di materiale (Guerrini 2020, p. 143). Essendo uno standard generale ISBD(G) non definì l'area 8 come *area dell'ISBN* o *area dell'ISSN*, ma utilizzò la denominazione: *area del numero standard e delle condizioni di disponibilità*, intendendo con numero standard qualsiasi tipo di identificatore che poteva comparire sulla risorsa. La locuzione *numero standard* continuò a essere utilizzata fino alla pubblicazione di *ISBD Edizione consolidata* del 2011, in cui la denominazione dell'area 8 si modificò significativamente in *Area dell'identificatore della risorsa e delle condizioni di disponibilità*, con la dichiarazione esplicita della funzione dell'identificatore; l'area divenne obbligatoria, aspetto che dimostra ulteriormente l'importanza di questo tipo di informazione.

Anche le AACR2, *Anglo American Cataloguing Rules 2nd Edition* pubblicate nel 1978 e riviste nel 2002 e nel 2005, erano strutturate seguendo il modello ad aree di ISBD e prevedevano un'area specifica per la registrazione del *numero standard*.

Il concetto di identificatore in ambito catalografico, tuttavia, fece la sua comparsa consapevole tra la fine degli anni Novanta del Novecento e i primi anni Duemila all'interno di FRBR, *Functional Requirements for Bibliographic Records* e degli altri modelli della famiglia *Functional Requirements* (FR). FRBR venne pubblicato nel 1998 e considerava gli identificatori importanti attributi per l'identificazione della manifestazione e dell'item. FRAD, *Functional Requirements for Authority Data*, pubblicato nel 2009, considerava l'*identificatore* come un'entità che, analogamente al *nome*, può essere associato a un'altra entità allo scopo di differenziarla da tutte le altre appartenenti a un determinato dominio. A differenza di FRBR, quindi, non concepiva più l'identificatore come un semplice attributo, ma come un'entità che poteva essere collegata a entità bibliografiche, ma anche a persone, famiglie ed enti.

Con la pubblicazione di FSRAD, *Functional Requirements for Subject Authority Data*, nel 2010, l'identificatore venne considerato come una sottoclasse dell'entità *Nomen* che rappresenta una designazione costituita una sequenza di segni utilizzata per riferirsi a un *Thema*, ossia "ogni entità utilizzata come soggetto di un'opera" (FRSAD 2010, p. 15).

Lo stesso anno venne pubblicata la prima edizione delle linee guida RDA, *Resource Description and Access* basate sui modelli concettuali della famiglia FR e in particolare FRBR e FRAD. Come FRBR, RDA considera gli identificatori degli attributi. Essi però non vengono registrati solo per la manifestazione e l'item, ma per tutte le entità descritte all'interno dello standard e quindi anche l'opera, l'espressione, la persona, la famiglia e l'ente. La prima edizione di RDA, inoltre, considera l'identificatore come un *elemento essenziale* della descrizione riconoscendo l'importanza che esso ha acquisito all'interno del contesto del web semantico. Gli elementi essenziali, infatti, devono sempre essere presenti al fine di considerare una descrizione conforme allo standard. Essi sono stati individuati in base alla loro capacità di svolgere le principali funzioni utente corrispondenti a:

- identificare e scegliere una manifestazione;
- identificare e scegliere le opere e le espressioni contenute in una manifestazione;
- identificare il creatore di un'opera;
- trovare una persona, un ente o una famiglia associati a una risorsa.

Nel 2017 viene pubblicato il nuovo modello concettuale IFLA *Library Reference Model* (LRM) che integra e sostituisce i modelli precedenti. Analogamente a FRAD e FRSAD considera gli identificatori come delle entità e, in particolare, come dei *nomen*:

un identificatore è un tipo di *nomen* destinato a essere persistente e unico nell'ambito di uno specifico dominio di applicazione, quali gli identificatori di pubblicazioni di tipo specifico o gli identificatori delle *persone*, in modo che le istanze dell'entità si possano specificamente identificare e vi si possa fare riferimento senza ambiguità. Ciò che distingue un identificatore da altri *nomen* è che il valore dell'attributo *stringa del nomen* di un identificatore non può essere identico al valore dell'attributo *stringa del nomen* di qualsiasi altro *nomen* all'interno di un determinato sistema (naturalmente, altri *nomen* al di fuori di quel sistema potrebbero avere lo stesso valore dell'attributo *stringa del nomen*). Gli identificatori sono generalmente assegnati in base a regole stabilite da agenzie autorizzate quali, ad esempio, le agenzie di registrazione di registrazione degli identificatori ISO e gli uffici della pubblica amministrazione per l'identificazione univoca di cittadini e residenti. L'ambito di un sistema basato su identificatori può essere generico (come per l'URI) o altamente specializzato (numeri di catalogo delle opere di un compositore) (IFLA 2020, p. 88).

In IFLA LRM, dunque, gli identificatori sono una tipologia specifica di *nomen* che si caratterizza per l'uso di una *stringa del nomen* univoca all'interno di un determinato contesto. La *stringa del nomen* consiste nella "combinazione di segni che forma una denominazione associata a un'entità tramite il *nomen*" (IFLA 2020, p. 53). Per esempio, la stringa di cifre 20150601 può costituire un attributo *stringa del nomen* di un ISSN. Il collegamento tra la *stringa del nomen* e l'entità a cui si riferisce da origine al *nomen* che, quindi, si presenta contemporaneamente come un'entità e come una relazione reificata tra entità.

IFLA LRM ha determinato una profonda revisione delle linee guida RDA che sono state modificate per allinearsi al nuovo modello concettuale. La seconda edizione dello standard è stata pubblicata nel dicembre 2020. Le trasformazioni introdotte hanno influito sul modo di registrare gli identificatori che non costituiscono più degli attributi per le entità ma dei *nomen* associati a esse mediante relazioni di denominazione. Il seguente schema mostra come può essere registrata la relazione tra un'entità persona e l'identificatore a essa associato.

The image shows a screenshot of a library system interface displaying an IRI (Internationalized Resource Identifier) element reference. The interface is light blue and contains the following information:

- Element Reference**: <http://rdaregistry.info/Elements/a/P50094>
- Domain**: Person →
- Range**: Nomen →
- Alternate labels**: has identifier for person
- IFLA LRM** (with a plus sign)
- MARC 21 Authority** (with a plus sign)

L'entità persona rappresenta il dominio da cui parte la relazione 'ha identificatore per persona', mentre il *nomen*, in questo caso un identificatore, rappresenta il codominio, ossia il punto di arrivo della relazione.

La normativa catalografica italiana al pari di quella internazionale riuscì a cogliere immediatamente l'importanza degli identificatori nella descrizione bibliografica. Le RICA, *Regole italiane di catalogazione per autori* del 1979 prevedevano la registrazione dell'ISBN o dell'ISSN nell'area delle note. Le successive REICAT, *Regole italiane di catalogazione* del 2009, affermano nell'introduzione che "devono costituire elementi di accesso i numeri e i codici identificativi attribuiti alle pubblicazioni e alle entità (p. es. a opere, a esemplari, a biblioteche, etc.) sia quando sono riconosciuti da standard internazionali o nazionali sia quando sono comunque largamente utilizzati. Quando esistono più codici, anche errati, la ricerca deve essere assicurata da tutti" (REICAT 2009, 0.4.2). A differenza di RDA in cui ogni identificatore è associato alla specifica entità a cui si riferisce REICAT concentrano la trattazione degli identificatori nel capitolo 4.8 *Area dei numeri identificativi*. Il paragrafo 4.8.1 è dedicato alla registrazione degli identificatori associati alla pubblicazione e, in particolare, ai numeri standard definiti dall'ISO quali: ISBN, ISSN, ISMN, ISRN, ISRC e ISAN. Il paragrafo 4.8.2 si concentra sui numeri e i codici attribuiti dall'editore, dal distributore o dal tipografo per identificare univocamente la pubblicazione o una delle sue parti. Infine, il paragrafo 4.8.3 riguarda le pubblicazioni antiche per le quali viene considerato come elemento identificativo l'impronta. Essa consiste in quattro gruppi formati da quattro caratteri ciascuno rilevati in punti predeterminati della pubblicazione seguiti da un numero o da una lettera che indica la fonte del terzo e del quarto gruppo di caratteri o dalla data di pubblicazione e da una lettera che indica come la data stessa si presenta.

Identificatori e formati di catalogazione

Nonostante i modelli concettuali e gli standard catalografici abbiano compreso già da tempo l'importanza degli identificatori "ben più complessa è la loro applicazione a flussi e pratiche catalografiche tradizionali, ancora fortemente vincolate alla strutturazione del record e all'utilizzo del formato MARC [MACHINE REABABLE CATALOGING]" (Moi 2021, p. 374). In questa fase di transizione, in attesa che i dati bibliografici vengano definitivamente prodotti in linked data, un ruolo fondamentale è svolto dal Program for Cooperative Cataloging che ha avviato una serie di sperimentazioni guidate dal PCC Task Group on URIs in MARC. Il Gruppo di lavoro si è posto l'obiettivo di definire delle modalità di arricchimento dei record MARC attraverso l'inserimento di URI autorevoli provenienti da fonti esterne al fine di garantire nel prossimo futuro una transizione dal MARC ai linked data più agevole possibile. Il MARC21 prevedeva già da tempo campi specifici dedicati alla registrazione di identificatori quali ISBN (020), ISSN (022) e altri numeri standard come l'ISTC e l'ISMN (024). Il sottocampo \$0, *Authority record control number or standard number* viene utilizzato dal 2007 per la registrazione del numero di controllo associato all'entità descritta o un URI contenente il numero di controllo come token (Hieh; Reese 2015, p. 212). Per esempio, nel record:

- 100 1# \$a Stravinsky, Igor,
- \$d 1882-1971,
- \$e composer.
- \$4<http://id.loc.gov/vocabulary/relators/cmp>
- \$0 <http://id.loc.gov/authorities/names/n79070061>

\$0 rappresenta la fonte da cui è stata ripresa la forma autorizzata del nome dell'entità persona registrata nel sottocampo \$a.

Nel 2019 il PCC Task Group on URIs in MARC ha approvato il sottocampo \$1 che è andato ad arricchire numerosi campi MARC con lo scopo di accogliere i *real world object (RWO) URIs*, ossia degli URI che identificano l'entità in se come oggetto del mondo reale. Per esempio:

- 257 ## \$a Korea (South)
- \$2 naf
- \$0 <http://id.loc.gov/authorities/names/n79126802>
- \$1 <http://vocab.getty.edu/tgn/7000299-place>

Il sottocampo \$1 contiene un URI che rimanda direttamente all'entità luogo riportata nel sottocampo \$a, mentre l'URI registrato in \$0 fa riferimento alla fonte autorevole da cui è tratta la forma del nome dell'entità.

La distinzione tra \$1 e \$0 consente di differenziare gli oggetti del mondo reale dalla loro descrizione. Per evitare confusione, infatti, è necessario utilizzare URI differenti in modo che l'oggetto e la descrizione siano chiaramente identificati. Per l'occhio umano è impossibile riconoscere la differenza tra un URI che rappresenta un RWO e un URI che rimanda alla sua descrizione. Per poterli distinguere è necessario utilizzare un programma che renda evidente lo scambio di messaggi tra client e server come Web Sniffer,¹ oppure è possibile utilizzare due strategie riportate nel documento *Cool URIs for the semantic web* del World Wide Web Consortium (W3C):

- il codice di stato *303 See Other*
- *Hash URI*

Il primo sistema utilizza il meccanismo della redirectione per restituire i dati RDF come descrizione della risorsa. A un client interessato a un oggetto del mondo reale il server risponde con un codice *HTTP 303 See Other* e con l'URI di un documento web che descrive l'oggetto richiesto. Per esempio, se il client chiede <https://www.wikidata.org/entity/Q200092> il server risponde *303 See Other* e invia al client l'identificatore URI relativo alla rappresentazione della risorsa, ovvero <https://www.wikidata.org/wiki/Special:EntityData/Q200092>. In questo modo il server dichiara che <https://www.wikidata.org/entity/Q200092> è un URI relativo a un RWO e che <https://www.wikidata.org/wiki/Special:EntityData/Q200092> è un URI relativo alla sua rappresentazione. Se il client avesse richiesto <https://www.wikidata.org/wiki/Special:EntityData/Q200092> il server avrebbe risposto *200 ok* comunicando così al client che la sua richiesta era stata ricevuta con successo e veniva fornito in risposta il contenuto richiesto. Quindi, il browser che usiamo quotidianamente quando incappa in un codice di stato 303 in maniera trasparente invia all'utente il contenuto dereferenziato del secondo indirizzo URI.

Il sistema *Hash URI* consiste nell'utilizzo di URI che contengono un frammento ossia una parte relativa alla specifica risorsa descritta separata dal resto dell'URI mediante il simbolo *hash* (#), per

¹ <https://websniffer.cc>

esempio, https://en.wikipedia.org/wiki/Bonnie_and_Clyde#Clyde_Barrow o https://en.wikipedia.org/wiki/Bonnie_and_Clyde#Bonnie_Parker. Quando il client vuole recuperare un *Hash URI* il protocollo http prevede che la parte preceduta dal simbolo # sia rimossa prima di chiedere l'URI al server (in questo caso verranno rimosse le parti #Clyde_Barrow e #Bonnie_Parker). Questo significa che un URI che include un *Hash* non può essere recuperato direttamente, ma vengono restituite tutte le descrizioni delle risorse appartenenti a uno specifico dominio indicato nell'URI stesso (https://en.wikipedia.org/wiki/Bonnie_and_Clyde).

La scelta tra i due sistemi non può dipendere da fattori oggettivi quanto piuttosto da considerazioni di opportunità dipendenti da specifici contesti di utilizzo: “il limite più rilevante dell'Hash URI è la quantità di dati che l'operazione di dereferenziazione restituisce. Infatti, di solito questa tecnica prevede la restituzione di un documento che contiene, al suo interno, una porzione (la parte locale o relativa) di informazione che corrisponde all'URI dereferenziato. È preferibile utilizzare URI 303 nel caso d'interrogazione di dataset corposi e utilizzare Hash URI per l'interrogazione di dataset poco consistenti” (Guerrini, Possemato 2015, p. 42).

Accanto all'arricchimento dei record MARC attraverso gli URI, nel corso degli ultimi anni sono stati sviluppati altri progetti per favorire il passaggio verso una descrizione interamente basata sui linked data. In particolare, va ricordata la Bibliographic Framework Initiative (BIBFRAME) che ha preso avvio nel 2011 e ha definito tra i suoi obiettivi principali quello di consentire il riutilizzo di milioni di record MARC prodotti dalle biblioteche nel contesto del web semantico attraverso i linked data e favorendo il riassetto dei dati delle attuali registrazioni MARC in nuove architetture coerenti che consentano una catalogazione collettiva a livello più granulare. Nel novembre 2012 la Library of Congress ha rilasciato la versione 1.0 di BIBFRAME, successivamente aggiornato nel 2016 con la pubblicazione della versione 2.0. BIBFRAME consiste di classi e proprietà RDF. Le classi principali sono *opera*, *istanza* e *item* a cui si aggiungono altre classi e sottoclassi. Le proprietà descrivono le caratteristiche della risorsa analizzata e le relazioni con altre risorse. Il nuovo modello attribuisce agli identificatori un ruolo centrale poiché riconosce che solo attraverso l'identificazione delle risorse mediante identificatori stabili, univoci e comprensibili da una macchina è possibile garantire la partecipazione delle biblioteche alla costruzione del web dei dati. Inoltre, il loro uso favorisce la condivisione dei dati tra le biblioteche riducendo i tempi necessari al lavoro d'autorità che ha sempre costituito uno degli aspetti più costosi della descrizione bibliografica:

BIBFRAME's use of controlled identifiers over the MARC format's reliance on controlled text strings for entity description will lessen considerably the time and costs associated with maintenance of authority data. One controlled string of data may appear in thousand of MARC records. If that controlled string of data changes, a fairly common occurrence, all MARC records containing that controlled string of data need to be changed as well. Maintenance of MARC records can be costly and time-consuming. A controlled identifier does not change, even if the controlled string of data associated with that identifier changes; BIBFRAME thus reduces the time and costs of bibliographic maintenance.²

² <https://www.librarianshipstudies.com/2017/12/bibframe.html>

Inoltre, l'ampio uso degli identificatori all'interno di BIBFRAME aumenterà la visibilità delle collezioni delle biblioteche di tutto il mondo. Essi, a differenza delle stringhe di testo adottate nel MARC non dipendono dalla lingua e ciò facilita la più ampia condivisione dei dati bibliografici. Il trattamento degli identificatori all'interno di BIBFRAME è stato modificato con il passaggio dalla versione 1.0 alla 2.0. Mentre la versione 1.0 definiva un'unica classe di identificatori considerando le singole tipologie di identificatore come proprietà, la versione 2.0 definisce un'unica proprietà per gli identificatori "bf:identifiedBy" e considera le diverse tipologie di identificatore come delle sottoclassi della classe principale identificatore. Quindi tipologie specifiche come ISBN, ISAN, DOI, LCCN ecc. sono delle sottoclassi di "identificarore" (bf:Isbn; bf:Isan; bf:Doi; bf:Lccn). La rappresentazione di queste tipologie come classi piuttosto che proprietà determina dei vantaggi in termini di riusabilità ed efficienza della ricerca (rende i dati più facilmente interrogabili).

Gli identificatori all'interno dei modelli e degli standard archivistici e museali

Accanto alle biblioteche anche gli archivi e i musei dimostrano di aver colto l'importanza degli identificatori per lo scambio e il riuso delle informazioni dedicando particolare attenzione al loro trattamento all'interno di standard e modelli concettuali. In ambito archivistico i principali standard adottati a livello internazionale come ISAD(G), *General International Standard Archival Description*; ISAAR(CPF), *International Standard Archival Authority Records for Corporate Bodies, Persons and Families*; ISDF, *International Standard for Describing Function* e ISDIAH, *International Standard for Describing Institution with Archival Holdings*, elaborati tra la fine degli anni Ottanta del secolo scorso e i primi anni Duemila prevedevano specifiche aree della descrizione destinate ad accogliere gli identificatori dell'entità descritta. ISAD(G), per esempio, consentiva di registrare l'identificatore all'interno dell'*Area dell'identificazione*. Esso poteva essere rappresentato dalla segnatura archivistica del documento o da un altro codice identificativo. Inoltre, lo standard precisava che l'identificatore insieme ad altri elementi della descrizione come la segnatura, la denominazione o titolo, il soggetto produttore, la data, la consistenza dell'unità di descrizione e il livello di descrizione era considerato essenziale per lo scambio di informazioni a livello internazionale. L'importanza degli identificatori nella descrizione archivistica è stata ulteriormente ribadita all'interno di RiC-CM, *Record in Context. A conceptual model for archival description*, un modello concettuale e un'ontologia per la descrizione archivistica elaborata dall'Expert Group on Archival Description (EGAD) tra il 2012 e il 2016. Gli identificatori vengono definiti come delle proprietà condivise da qualsiasi entità riconosciuta dal modello. In particolare, vengono distinte due categorie di identificatori:

- RiC-P1³ identificatori globali persistenti: sono degli identificatori globali unici e risolvibili in modo persistente per le entità e si presentano come degli URI, per esempio: <http://n2t.net/ark:/99166/w6v1266v> è un identificatore ARK associato a un record; <http://isni.org/0000000073572182> è un identificatore ISNI per una persona;
- RiC-P2 identificatori locali: sono degli identificatori assegnati localmente a un'entità, per esempio: BUD-01-F002 è un numero di classificazione assegnato a un file aziendale; NAS1/A/1.1 è un identificatore locale associato a un record.

³ Denominazione prevista da RiC-CM per le proprietà.

In ambito museale l'ontologia maggiormente diffusa a livello internazionale è CIDOC CRM, *Conceptual Reference Model* elaborata a partire dal 1994 dall'International Committee for Documentation (CIDOC) dell'International Council of Museums (ICOM) con l'obiettivo di fornire un modello di riferimento e uno standard che musei e altre istituzioni culturali possono utilizzare per descrivere le loro collezioni e favorire la condivisione delle informazioni. Il modello prevede delle classi principali che rappresentano le entità a cui sono associate delle proprietà che le descrivono e stabiliscono delle relazioni tra esse. Gli identificatori sono una sottoclasse della classe principale *Denominazione* (E41) e comprendono stringhe e codici assegnati alle istanze e alle entità previste dal modello per identificarle in modo univoco e permanente nel contesto di una o più organizzazioni. Essi possono coincidere con numeri di inventario, codici di registrazione ecc. Indirizzi postali, numeri di telefono, URL e indirizzi di posta elettronica sono altresì considerati identificatori utilizzati nei servizi di trasporto tra clienti (CIDOC CRM 2022, p. 87).

Conclusioni

Le diverse istituzioni della memoria registrata hanno saputo cogliere l'importanza degli identificatori per la descrizione e l'accesso alle risorse fin dalla loro comparsa. Inoltre, in seguito allo sviluppo del web e delle nuove tecnologie hanno messo in atto specifiche politiche di gestione al fine di sfruttare le loro potenzialità per lo scambio, la condivisione di dati e per il controllo d'autorità. Il peso attribuito agli identificatori si coglie, inoltre, dall'analisi dei modelli concettuali e degli standard prodotti in ambito biblioteconomico, archivistico e museale in cui gli identificatori vengono spesso considerati come elementi essenziali per la descrizione delle risorse sia che vengano considerati come attributi o proprietà di una risorsa, sia come sottoclassi di entità principali, per esempio, il *nomen* in IFLA LRM. In questo caso è necessario stabilire una relazione di denominazione tra l'identificatore e l'entità a cui si riferisce. Gli standard prevedono la possibilità di registrare sia identificatori globali supportati da standard ISO e diffusi a livello internazionale sia identificatori locali, ma riconoscono la superiorità dei primi per la condivisione e il riuso dei dati. Per arrivare a sfruttare a pieno le potenzialità degli identificatori, tuttavia, non è sufficiente riconoscerne l'importanza nell'ambito della descrizione bibliografica e documentale, ma necessario che i formati di catalogazione e metadatozione oggi a disposizione delle varie comunità si trasformino al fine di allinearsi ai linguaggi e alle tecnologie del web semantico.

Riferimenti bibliografici

Boulet, Vincent. 2022. "How to build an 'Identifier's policy': the BnF use case". *JLIS.it* 13(1): [177]-184. DOI: <https://doi.org/10.4403/jlis.it-12768>.

Commissione permanente per la revisione delle regole italiane, ed. 2009. *Regole italiane di catalogazione REICAT*. Roma: ICCU.

Concedda, Flavia. 2015. "ISSN e I suoi fratelli: la lunga marcia dei codici internazionali: prima parte: UNESCO, ISO e il primogenito". *AIB Studi* 55(2): 295-308.

Concedda, Flavia. 2015. "ISSN e i suoi fratelli: la lunga marcia dei codici internazionali: parte seconda: il CNR e la storia italiana dell'ISSN. Gli altri sistemi identificativi". *AIB Studi* 55(3): 439-457.

Guerrini, Mauro. 2020. *Dalla catalogazione alla metedatazione. Tracce di un percorso*. Roma: Associazione italiana biblioteche.

Guerrini, Mauro, Tiziana Possemato. 2015. *Linked data per biblioteche, archivi e musei*. Milano: Editrice Bibliografica.

Heih, Jackie and Terry Reese. 2015. "The importance of identifiers in the new web environment and using the Uniform Resource Identifier (URI) in subfield zero (\$0): a small step that is actually a big step". *Journal of Library Metadata* 15(3-4): 208-226. <https://doi.org/10.1080/19386389.2015.1099981>.

IFLA (International Federation of Library Association and Institutions). (2010). *Functional Requirements for Subject Authority Data (FRSAD). A conceptual model*. Edited by Marcia Lei Zeng, Maja Žumer, Athena Salaba. Berlin: De Gruyter.

IFLA (International Federation of Library Association and Institutions). 2020. *IFLA Library Reference Model. Un modello concettuale per le informazioni bibliografiche*. Edited by IFLA Functional Requirements for Bibliographic Records (FRBR) Review Grup, Pat Riva, Patrick Le Boeuf, Maja Žumer, italian edition edited by Istituto centrale per il catalogo unico (ICCU). Roma: ICCU.

Moi, Alessandra. 2021. "Da universale a globale. Il controllo bibliografico nel web semantico". In *La trasmissione della conoscenza registrata*, a cura di Carlo Bianchini e Lucia Sardo, 369-378. Milano: Editrice Bibliografica.

Vitiello, Giuseppe. (2004). "L'identificazione degli identificatori". *Biblioteche oggi* 2: 67-80.

W3C Interest Group. 2008. *Cool URI for the semantic web*. <https://www.w3.org/TR/cooluris/>.