

# Per un nuovo catalogo SBN e per una nuova Bibliografia nazionale italiana

**Mauro Guerrini**

*Università di Firenze*

## Gli scopi originari di SBN

**S**BN nasce con l'obiettivo di offrire tre fondamentali servizi al Paese: il controllo bibliografico e l'informazione bibliografica sulla produzione editoriale italiana corrente e retrospettiva; la disponibilità delle pubblicazioni tramite una rete di prestito interbibliotecario nazionale, aperta agli scambi con l'estero; la formazione professionale e tecnica dei bibliotecari che operano all'interno del Servizio bibliotecario italiano. Se esaminiamo SBN dal punto di vista del catalogo, ci sono tre contributi che nel corso degli anni si sono consolidati e costituiscono a oggi punti forti e predisposti a evolversi in futuro.

Il primo è costituito dalla cooperazione tra biblioteche di natura e dimensioni differenti con la partecipazione alla redazione e alla condivisione di notizie bibliografiche: ciò ha dato origine a un progetto volto alla produzione cooperativa di servizi e alla rappresentazione qualitativa della complessa realtà bibliografica nazionale, tramite un catalogo collettivo e unitario. Il secondo è dato dalla struttura informatica che nasce sin da subito granulare anziché chiusa nel concetto rigido di record bibliografico (pensiamo al reticolo tramite il quale si costruisce l'informazione, con VID e BID che identificano singoli elementi): vi è stata, cioè, l'intuizione dell'importanza del *dato in sé*, concetto di base che condurrà ai linked data odierni. Il terzo consiste nell'essere l'unica struttura bibliografica di estensione e rilevanza nazionale esistente al mondo (SBN è il riferimento imprescindibile per la ricerca di libro nelle biblioteche italiane), all'interno

della quale sono confluite biblioteche statali, di enti locali, delle università, scolastiche, ecclesiastiche, di accademie e istituzioni pubbliche e private operanti in diversi settori disciplinari; assistiamo in ciò a una forma di cooperazione tra Stato centrale, Università e Regioni. SBN è, pertanto, un contributo che la comunità dei bibliotecari offre al Paese.

SBN ha avuto dei limiti, come la natura chiusa dei linguaggi e degli applicativi utilizzati che, a fronte di un'idea innovativa di apertura a tutte le realtà, si è fatta nel tempo ingabbiare entro i confini stretti dello specifico software utilizzato e ha viziato il mercato nazionale dei software. Con l'evoluzione dell'Indice SBN e con il nuovo protocollo SBN MARC, tuttavia, numerosi sono stati gli elementi innovativi del sistema, tra i quali: l'apertura dell'Indice SBN a sistemi di gestione non SBN nativi che scambiano messaggi con il sistema centrale tramite il protocollo SBNMARC che utilizza la semantica MARC, gestito dall'ICCU; la gestione di livelli di cooperazione diversificati in cui ciascun polo può scegliere il proprio livello di partecipazione a SBN; lo sviluppo di funzionalità specifiche per la catalogazione di tutte le tipologie di risorsa.

Partendo da queste premesse, si può disegnare uno scenario futuro più coerente con le scelte che grandi biblioteche dell'Europa e del mondo stanno compiendo.

Il catalogo si trasformerà in un *portale* e rientrerà nella famiglia dei discovery tool; esso manterrà e aumenterà la sua forza di strumento affidabile ed efficace nella ricerca e nel recupero dell'informazione, nonché nelle moda-

lità di presentazione dei risultati, grazie alla tecnologia dei linked data che rende ogni dato autonomo e assemblabile in una molteplicità di forme diverse. Il nuovo strumento di mediazione e di fornitura diretta dei testi (*full text*) rispetterà le *funzioni utente* di FRBR – *trovare, identificare, selezionare, ottenere* – a cui gli ICP del 2009 hanno aggiunto la funzione di *navigare*. Non solo, esso potrà raggruppare i risultati secondo parametri differenti, aiutando così l'utente a *individuare, selezionare e ottenere* la risorsa che meglio risponda alla ricerca impostata. Consentirà, prima ancora, l'efficace visione a grappolo di record differenti ma con caratteristiche comuni: per esempio, la nuova espressione della medesima opera. L'attuale presenza di notizie bibliografiche per ciascuna differente edizione, traduzione, riduzione della medesima opera non è funzionale alla ricerca, perché costringe l'utente ad aprire numerosi record. Con la nuova prospettiva, invece, sarà possibile la visione panoramica dell'*œuvre* di un autore, cioè dell'unità della sua produzione letteraria (come ricordava Ákos Domanovszky).

Un altro aspetto decisivo sarà il superamento della filosofia del possesso anche per i dati e non solo per le raccolte, come ricordava Karen Coyle lo scorso anno al convegno *Global interoperability and linked data in libraries*, i cui atti sono stati pubblicati su J LIS.it vol 4, n. 1, 2013. Ci sono aspetti economici da non trascurare. I dati prodotti dagli editori sono l'esito di un ente profit mentre i dati delle biblioteche sono realizzati da istituzioni non profit. L'utilizzo libero dei dati da parte del web necessita pertanto di ingenti finanziamenti pubblici.

### Il catalogo come collaborazione

Nel web costituito dai linked data, non esiste più la distinzione tra dati bibliografici e altri dati: esistono solo dati condivisibili, modulari e riutilizzabili, indipendentemente da chi li ha creati.

L'esperienza della *Bibliografia nazionale svedese* indica che la strada sarà il trattamento delle

proprie risorse interne come se fossero esterne: il lavoro di authority si trasforma dalla produzione locale alla verifica della qualità e della coerenza dei dati prodotti da altre agenzie e al riutilizzo dei dati "controllati". L'agenzia bibliografica nazionale e, in generale, tutte le biblioteche, avranno il compito di individuare, tra le molte fonti disponibili, quelle più autorevoli, e collegare dati tra loro coerenti. La qualità è pertanto sinonimo di affidabilità.

La percentuale di riutilizzo, dell'interoperabilità dei dati, secondo alcuni calcoli, arriva fino all'80%.

Un vantaggio non trascurabile per l'ente che metta a disposizione i propri dati è che enti terzi possono costruirsi sopra progetti e servizi, indirizzati anche all'ente proprietario che riceve, così, un immediato beneficio.

Un interessante esempio di applicazione trasversale dei linked data è il progetto realizzato dal Comune di Firenze: i 230 data set creati per lo stradario sono stati incrociati con il tesoro del Nuovo Soggettario e con il VIAF per consentire di individuare vie e piazze dedicate a personaggi (e più in generale a entità) i cui nomi sono utilizzati nei rispettivi archivi. Analogo link, biunivoco, è con i termini di Wikipedia Italia: quello del *Nuovo Soggettario* è il primo tesoro universale a cui essa sia collegata. Su wikidata vi è stato anche un primo caricamento di record relativi alle voci di autorità degli autori presenti in SBN ed è stato attivato il link tra Wikipedia in italiano e l'OPAC SBN. Partendo da un nome "controllato", nel senso degli ICP, presente in Wikipedia in italiano, è quindi possibile consultare la relativa scheda di autorità dell'OPAC SBN e avviare la navigazione sulle risorse bibliografiche collegate.

Un'applicazione che esemplifica l'approccio dei linked data ai dati bibliografici è il progetto *Des fiches de référence sur les auteurs, les œuvres et les thèmes* (data.bnf.fr) concretizzato dalla Bibliothèque nationale de France. Il lettore è guidato alle risorse tramite una pagina unica che raggruppa tutte le informazioni provenienti dai diversi cataloghi e da Gallica.

*Des fiches de référence* supera il concetto tradizionale di catalogo e si presenta come un portale enciclopedico. Sulla stessa prospettiva si pongono progetti di molte altre biblioteche, dalla British Library, alla Nazionale spagnola, alla Deutsche Nationalbibliothek. La possibilità di navigare per autore, per opera o per soggetto lascia percepire come potrebbe essere costruito il catalogo del futuro, sempre più risultato di un lavoro sul creatore e sulla sua oeuvre. Tutto ciò presuppone la necessità di produrre soluzioni intersettoriali, cioè valide per le biblioteche, i musei e gli archivi.

Esisteranno ancora il record bibliografico e il catalogo? La registrazione cambierà radicalmente aspetto, nella direzione dell'atomizzazione; il nuovo record sarà costituito da un insieme di dati corredati di significato non ambiguo, significato che sarà espresso in un glossario depositato in un open repository con definizioni redatte e mantenute da esperti e da agenzie autorevoli. I dati, soprattutto, saranno connessi secondo rigorose regole *semantiche*. La registrazione, pertanto, sarà del tutto rivoluzionata: dalla sostanziale staticità che la contraddistingue adesso, essa diverrà la risultante dinamica di un processo di aggregazione di dati, la cui provenienza, per una parte sempre più consistente, sarà estranea al lavoro di una biblioteca o di un'agenzia bibliografica, e potrà essere generata da soggetti diversi, come editori, librai, enti di ricerca e organi amministrativi.

### **SBN nella nuvola semantica dei linked data**

Il nuovo modello dell'Indice SBN dovrebbe utilizzare la tecnologia degli open linked data in modo conforme ai principi e alle finalità del web semantico; i dati bibliografici dovrebbero essere pubblicati sul web in formati standard che ne garantiscano l'operabilità, ovvero in modalità leggibile, interpretabile e utilizzabile dalle macchine, affinché siano visibili, disponibili e liberamente riutilizzabili, come parte integrante dell'universo informativo.

RDA – che si presenta come nuovo standard internazionale – costituisce l'elaborazione più aderente a FRBR oggi disponibile e presenta una rinnovata enfasi sui punti d'accesso dopo decenni di attenzione alla descrizione; suo obiettivo principale è la collaborazione con le comunità di produttori di metadati dell'intero web. Ciò rende il codice pronto a essere utilizzato nel nuovo scenario dei linked data; per questo motivo la Germania e la Francia hanno deciso di divenire soggetti protagonisti nella sua redazione. Le RDA saranno impiegate da numerosi paesi europei (EURIG) e del mondo, Cina compresa.

Si apre un insieme di questioni che, prima o poi, SBN dovrà affrontare se sbarcherà in modo definitivo nel web semantico migrando verso il cloud computing. La risorsa catalografica dovrà difatti prevedere l'adattabilità a tutti i contenuti esistenti in rete, l'adattabilità a tutti i media possibili e un grado di modularità, tale da accogliere le risorse innovative e in parte imprevedibili del futuro, nonché un sistema di archivi di dati autonomi e interoperabili (tra essi e con risorse esterne) per ciascun gruppo di entità o per ciascuna entità prevista dagli standard FRBR e FRAD.

Più critica risulta la questione della esternalizzazione di dati e software: l'adesione alla modalità cloud computing, che sembra essere destinata a prevalere su scala globale per l'ineludibile questione del contenimento dei costi e dell'acquisizione di vantaggio nel contesto digitale, dove la competizione è trasversale e si regge sulla popolarità delle risorse tra i naviganti, comporta che l'operatore che ricorre alla tecnologia informatica agisca tendenzialmente da remoto e on-demand su infrastrutture esternalizzate con varie modalità e con costi limitati al loro effettivo utilizzo. Nelle forme più spinte, l'operatore non è coinvolto né nella gestione delle componenti software e hardware né nello stoccaggio dei dati. Bypassando gli investimenti diretti in tecnologia aggiornata costantemente e rapidamente destinata all'obsolescenza (perciò

poco conveniente da acquisire in prima persona) e in personale continuamente formato per valutarla e gestirla, l'operatore può concentrarsi esclusivamente sulla propria missione organizzativa e gestire sotto forma di costi operativi (non più, dunque, strutturali) il supporto tecnologico necessario. Se la tendenza della rete sarà questa, è evidente che le organizzazioni pubbliche e private in essa operanti non potranno fare altro che adeguarsi.

Le tecnologie sono in pieno divenire e ciò comporta una serie di specifiche criticità, le quali, in sintesi, sono:

- la difficile distinzione tra servizi pienamente cloud e hosting tradizionale spacciato per cloud;
- la totale dipendenza dei servizi dalla rete nelle cloud (anche solo a gestione pubblica) di grande portata;
- l'assenza di standard tecnologici e di rilevamento statistico (metrics e bibliometria);
- la sicurezza dei dati, che può essere mitigata distribuendo le risorse in modalità ibrida tra pubblico e privato;
- la gestione e la riservatezza dei dati sensibili, affrontabile come la criticità precedente;
- la durabilità e la preservazione dei dati, affrontabile come la criticità precedente, ma è legittimo chiedersi fino a che punto;
- la regolamentazione giuridica della tecnologia e del diritto d'autore frammentata tra stati nazionali;
- la questione della estensibilità del diritto d'autore dalla creazione dei dati, date le caratteristiche tecnologiche del web semantico e della nuvola, al loro assemblaggio in una forma ritenuta originale;
- l'affidabilità tecnologica e la correttezza etica dei gestori delle nuvole, il controllo delle loro attività nonché degli usi o ri-usi (più o meno noti o consentiti) che possono fare dei dati (si noti che il riutilizzo dei dati è connaturato al web semantico e ciò costituisce un grosso problema sotto diversi punti di vista);

- la capacità di distinguere tra politiche a breve termine che mirano a conseguire risparmi immediati (e sono dunque avvertite come priorità) e politiche a lungo termine che avendo una valenza strategica e culturale risultano quasi sempre diseconomiche nell'immediato (quindi scarsamente allentanti nei confronti dei cosiddetti stakeholders delle biblioteche: comunità, decisori politici, opinion leaders).

Potrebbe, inoltre, aggiungersi presto un problema specifico. Se l'interfaccia a campo unico e generale di ricerca è già una realtà e la registrazione bibliografica lineare e interpretabile nel contesto testuale denominata MARC sta per morire o è già morta (per riprendere un famoso articolo di Roy Tennant), dal momento che tale contesto non è migrato sul web, allora la questione cruciale (come fa notare Karen Coyle) non sarebbe tanto quella della conversione in linked data degli archivi esistenti bensì quella di una forma del dato e del sistema d'accesso al dato che risulti *in partenza* e in modo nativo interoperabile col web semantico. In altre parole le ricerche del futuro prescinderebbero dal ricorso a risorse web non semantiche (banche dati, opac o repertori bibliografici redatti in modalità testuale e non costituiti da linkable-data), col rischio che le risorse tradizionali, non solo quelle catalografiche ma, a questo punto, anche quelle documentarie, finiscano per collocarsi in un qualche strato sommerso o del tutto esterno alla rete, senza che ci sia la possibilità di farle ri-emergere in superficie per esporle alla fruizione dell'utenza.

### La nuova BNI

Va compiuta una riflessione sul rapporto tra SBN e la *Bibliografia nazionale italiana*, un servizio indispensabile, la cui funzione non è purtroppo sempre compresa, perfino dagli stessi bibliotecari. Il controllo bibliografico nazionale autorevole, tempestivo ed esaustivo è un servizio indispensabile per SBN e per tutte

le biblioteche italiane e del mondo; deve, però, essere efficiente. Senza il censimento della produzione di cultura, l'Italia editoriale scompare dal circuito informativo internazionale: così viene meno un aspetto legato alla stessa identità culturale del nostro Paese. Con un controllo bibliografico efficiente, inoltre e soprattutto, molte biblioteche potrebbero dedicarsi più utilmente a potenziare i servizi all'utenza, avvalendosi delle notizie bibliografiche prodotte da un'agenzia specializzata, con un risparmio economico in termini generali per l'Italia, ma soprattutto con la disponibilità di un servizio di qualità. Ciò, oggi, avviene solo in minima parte; il servizio dovrebbe essere notevolmente potenziato dalla realizzazione di una nuova bibliografia italiana, sperimentando modelli di collaborazione con agenzie private di comprovata qualità, come Casalini Libri, in un progetto comune che garantisca un vero completo controllo della produzione italiana corrente, che ammonta, oggi, a circa 100.000 risorse bibliografiche l'anno.

Come sappiamo, da circa 50 anni in Italia esistono due bibliografie: la *Bibliografia nazionale italiana*, della BNCF, e *I libri* della Casalini Libri (peraltro la prima redatta a Firenze e la seconda a Fiesole). Casalini partecipa al programma cooperativo americano PCC (Program for Cooperative Cataloging) e condivide i requisiti di qualità della Library of Congress. Casalini, inoltre, crea e mantiene le registrazioni di autorità in NACO (Name Authority Cooperative Program of the PCC) per nomi, enti, convegni, luoghi geografici e titoli, e SA-CO (Subject Authority Cooperative Program of the PCC) per i soggetti: quale occasione migliore per partecipare attivamente all'importante progetto VIAF, a cui l'Italia, tramite l'ICCU, ha purtroppo potuto dare finora un apporto limitato (finora solo 45 mila su un totale di tre milioni di autori in SBN – a fronte di sette milioni in Germania – di cui solo 1300 sono su Wikidata e quindi su Wikipedia). La BNI censisce attualmente circa 6 mila risorse l'anno (di cui il 10% di traduzioni), mentre

Casalini arriva a oltre 21.000 record. La sovrapposizione, per buona parte, delle due iniziative – che non ha corrispondenze al mondo – non ha alcun senso; la ragionevolezza e la buona gestione delle cose invitano a individuare una soluzione condivisa, nell'interesse del Paese, prendendo a modello la fruttuosa collaborazione tra Library of Congress e il consorzio OCLC. Ciò dovrebbe avvenire assieme alla riflessione sui compiti delle bibliografie nazionali nell'era digitale, come stabilito dalle linee guida dell'IFLA emanate nel 2009, che includono nella bibliografia nazionale il censimento delle risorse elettroniche (ebook, periodici online e, in generale, le risorse digitali) e prevedono forme ampiamente collaborative tra istituti diversi che partecipano alla sua redazione.

## Proposte

L'evoluzione di SBN verso i linked data non dovrebbe essere l'ennesimo progetto volto a sviluppare un servizio o un sistema parallelo. È necessario un ripensamento complessivo e profondo di SBN, del suo data base e della sua architettura, per implementare un sistema in grado di produrre, gestire, condividere ed esporre open linked data, usando tecnologie più flessibili, secondo i fabbisogni della *mobile technology*.

Occorre che SBN:

- conti su una direzione scientifica efficiente, anziché solo su comitati di gestione, che hanno spesso anteposto gli interessi degli enti rappresentati al fine di redistribuire a pioggia i finanziamenti disponibili, senza mai elaborare una strategia di medio-lungo periodo;
- disponga di una propria divisione informatica: da trent'anni l'istituto proprietario delle attrezzature fa sviluppare e gestire i servizi a terze parti; ciò ha generato una forte dipendenza e la percezione di un mondo chiuso, controllato da pochi addetti ai lavori;
- disponga di nuovo personale competente

- e motivato al quale il personale dell'ICCU, ridotto come numero, possa trasmettere le proprie competenze e le informazioni per garantire la futura gestione delle attività;
- investa in innovazione;
  - prosegua la collaborazione con i progetti e le iniziative internazionali, soprattutto per gli standard;
  - usi un codice di regole affidabile e flessibile, condiviso a livello internazionale;
  - accetti software commerciali dialoganti con l'Indice, anziché solo SBN nativi; le esperienze delle Università di Padova e di Firenze sono molto importanti; al momento i software commerciali devono chiedere la certificazione (dietro pagamento) all'ICCU;
  - affermi sempre più chiaramente che esso è un sistema distribuito di risorse – umane e professionali – e riconosca che record di qualità possono essere prodotti da vari soggetti, valorizzando le competenze distribuite in tutto il territorio italiano;
  - introduca un cambiamento in termini di servizi, riprendendo e rilanciando il prestito interbibliotecario, mai realmente promosso, come risultava dal progetto iniziale;
  - renda sempre più fruttuoso il rapporto con una nuova *Bibliografia nazionale italiana*, un servizio essenziale, costruito sull'idea di una partnership pubblico-privato (BNCF-Casalini libri), in un quadro di cooperazione fra biblioteche, finalizzata alla sua produzione cooperativa, seguendo le linee guida IFLA sul tema del controllo bibliografico nell'era digitale.

Il catalogo SBN è un riferimento per i bibliotecari e per gli utenti italiani e stranieri da quarant'anni: non può morire per decisione amministrativa. Esso, però, dev'essere gestito in modo aperto ed efficiente da diventare attrattivo; perciò SBN dev'essere rilanciato, e quindi adeguatamente finanziato, per essere davvero il perno del servizio bibliotecario nazionale.

L'ultima consultazione dei siti web è avvenuta nel mese di dicembre 2013.