

Rivista del digitale nei beni culturali



ICCU

Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche https://www.iccu.sbn.it

Copyright © ICCU - Roma

La riproduzione totale o parziale del contenuto della rivista è ammessa con obbligo di citazione

«Dig*Italia*», rivista del digitale nei beni culturali, è una rivista *peer reviewed*.

I contributi possono essere proposti alla rivista dagli autori o su iniziativa del Comitato Scientifico.

Gli articoli vengono sottoposti al Comitato di Redazione per un primo accertamento sulla corrispondenza con i campi di ricerca della rivista. I contributi delle sezioni Saggi e Progetti vengono indirizzati in forma anonima ad almeno uno studioso di comprovata competenza sui temi affrontati. I revisori fanno pervenire i loro giudizi alla redazione (favorevole alla pubblicazione, favorevole con modifiche/miglioramenti, non favorevole). Se il giudizio finale è positivo, viene comunicata agli autori l'accettazione del contributo, insieme ad eventuali indicazioni suggerite dai valutatori, di cui si garantisce comunque l'anonimato.



Rivista del digitale nei beni culturali ISSN 1972-6201 Anno XIV. Numero 2 - Dicembre 2019

In copertina:

L'immagine è una libera elaborazione grafica della testa della statua di Apollo del I sec. d.c. (Civitavecchia, Museo Nazionale), copia da un originale greco avvicinabile all'Apollo di Leochares (IV sec. a.c.)

Direttore Fondatore

Marco Paoli

Direttore Responsabile

Simonetta Buttò

Comitato di Redazione

Capo Redattore: Elisabetta Caldelli

Amalia Maria Amendola Valentina Atturo Lucia Basile Laura Borsi Flavia Bruni Elisabetta Castro Massimina Cattari Silvana de Capua Carla Di Loreto Maria Cristina Di Martino Vilma Gidaro Egidio Incelli Maria Cristina Mataloni Massimo Menna Lucia Negrini Paola Puglisi Alice Semboloni Vittoria Tola Maria Lucia Violo

Grafica & Impaginazione

MLA&Partner - Roberta Micchi

Produzione e Stampa

Istituto Poligrafico e Zecca dello Stato S.p.A. Roma

Editore

ICCU
Istituto centrale per il catalogo unico delle biblioteche italiane
e per le informazioni bibliografiche
Viale Castro Pretorio, 105
00185 Roma
T +39 06 49.210.425
F +39 06 49.59.302
email: digitalia@iccu.sbn.it
http://digitalia.sbn.it

In attesa di registrazione al Tribunale di Roma



Comitato Scientifico

Osvaldo Avallone Giovanni Bergamin Dimitri Brunetti Simonetta Buttò Rossella Caffo Rosaria Campioni

Maria Carla Cavagnis Sotgiu

Laura Ciancio Flavia Cristiano Gianfranco Crupi Andrea De Pasquale Maria Cristina Di Martino

Pierluigi Feliciati Marina Giannetto Maria Guercio Mauro Guerrini Klaus Kempf Patrizia Martini Maurizio Messina Maria Cristina Misiti Maria Teresa Natale Marco Paoli

Don Valerio Pennasso Alberto Petrucciani Massimo Pistacchi Marco Pizzo Paola Puglisi Roberto Raieli Gino Roncaglia

Giovanni Solimine Laura Tallandini Anna Maria Tammaro Costantino Thanos Antonella Trombone Paul Gabriele Weston

Maria Letizia Sebastiani

SOMMARIO dicembre 2019

SAGGI	
Le opere orfane di Franca Berioli	9
Digitalizzare i manoscritti medievali? Riflessioni a margine di un progetto internazionale di Francesco Siri	33
PROGETTI	
Biblioteche in cammino L'esperienza del CoBiS: i progetti Linked Open Data e Digital library di Elena Borgi, Valeria Calabrese, Gabriella Morabito, Emanuela Secinaro, Anna Maria Viotto	43
Istituti culturali ecclesiastici: una visione generativa, inclusiva, sostenibile e prospettica di Francesca Maria D'Agnelli, Silvia Gallarato, Maria Teresa Rizzo	62
Il progetto di digitalizzazione del patrimonio storico della Biblioteca di Area Agraria dell'Università di Napoli "Federico II" di Nicola Madonna	87
Dai servizi digitali alla terza missione. L'evoluzione dei servizi bibliografici digitali all'Università degli Studi di Torino e la nascita del Polo bibliografico della ricerca di Oriana Bozzarelli, Franco Bungaro, Viviana Mandrile, Elena Marangoni, Maria Vittoria Muzzupapa	101

116	La memoria nel tempo delle reti e del digitale L'Archivio storico della Presidenza della Repubblica e il suo Portale storico di Marina Giannetto
138	Sfogliare gli album fotografici dell'UDI per ripercorrere la storia dell'Italia di Letizia Leo, Ilaria Scalmani, Vittoria Tola
	DOCUMENTI E DISCUSSIONI
161	I nuovi e principali Programmi di finanziamento europei (2021-2027). Proposte della Commissione Europea e documentazione di Massimina Cattari, Veronica Carrino
169	L'inaugurazione della Biblioteca Digitale della Biblioteca nazionale centrale di Roma Roma, BNCR, 12 dicembre 2018 di Fabio D'Orsogna
177	SEGNALAZIONI Violenza maschile contro le donne: digitalizzare una lunga storia di Vittoria Tola

Biblioteche in cammino L'esperienza del CoBiS: i progetti Linked Open Data e Digital library

Elena Borgi - Accademia delle Scienze di Torino
Valeria Calabrese - Biblioteca civica Centrale di Torino
Gabriella Morabito - Deputazione Subalpina di storia patria
Emanuela Secinaro - Istituto Nazionale di Ricerca Metrologica (INRIM)
Anna Maria Viotto - Associazione Archivio Storico Olivetti

L'esperienza LOD (linked open data) del CoBiS comprende al momento 12 biblioteche che fanno parte del Coordinamento biblioteche speciali e specialistiche dell'area metropolitana torinese. Il progetto, finanziato dalla Regione Piemonte, prevede la condivisione dei dati catalografici delle singole biblioteche sul portale https://dati.cobis.to.it/. Si è creato così un nuovo luogo online che grazie alla tecnologia dei linked open data consente all'utente di ricercare i dati catalografici ma anche di acquisire nuove informazioni grazie al collegamento ad altre fonti esterne come, ad esempio, wikipedia. La digitalizzazione di fondi bibliografici e archivistici di Istituti afferenti al CoBiS (CoBiS Digital Library & Archives) sarà prossimamente integrata al progetto CoBiS LOD. Un miglioramento del servizio per gli utenti nel panorama italiano e straniero delle sperimentazioni legate ai linked open data.

1. Il Coordinamento delle Biblioteche Speciali e Specialistiche dell'Area Metropolitana Torinese

I CoBiS, Coordinamento delle Biblioteche Speciali e Specialistiche dell'Area Metropolitana Torinese, nasce il 18 giugno 2008 su iniziativa di alcune biblioteche specialistiche invitate dalla Regione Piemonte a cercare un confronto stabile fra loro¹. L'obiettivo è ancora oggi quello di creare una rete tra biblioteche affini non per materia trattata ma per specificità di utenza e di servizi, biblioteche che non appartengono alla rete delle biblioteche civiche o delle biblioteche universitarie, ma che costituiscono per il Piemonte una importante realtà nel panorama culturale e conservano un considerevole patrimonio documentale.

A sollecitare i responsabili delle diverse biblioteche a creare un coordinamento fu Eugenio Pintore, dirigente dell'Assessorato alla Cultura – Settore Istituti culturali e biblioteche, prematuramente scomparso il 27 agosto 2019. A lui, prima bibliotecario e poi funzionario attento e sensibile, è dedicato questo contributo.

Le finalità del Coordinamento sono: promuovere una rete tra le differenti realtà documentarie a vantaggio di un proficuo servizio per le istituzioni e il pubblico a livello cittadino, regionale e nazionale; promuovere momenti di formazione continua per il personale specializzato afferente alle singole biblioteche; migliorare la visibilità e la fruibilità delle risorse documentali conservate, attraverso eventi e progetti.

Il CoBiS è oggi costituito da 66 biblioteche con varie specializzazioni, dall'area storica a quella artistica, da quella economico-sociale a quella scientifica². Le biblioteche sono afferenti a istituzioni private e pubbliche, di più recente costituzione o storiche, con patrimoni documentali assai diversificati, variabili da qualche migliaio di volumi a oltre 300.000.

Tutte le biblioteche del CoBiS sono infine aperte al pubblico e condividono una tipologia di utenza differente da quella delle biblioteche universitarie o di pubblica lettura: il personale addetto ha una particolare formazione, adatta al tipo di materiale conservato e alle richieste specifiche degli utenti.

Dal punto di vista organizzativo il CoBiS è caratterizzato da un ruolo attivo e costante dei bibliotecari partecipanti, che rappresentano la singola biblioteca aderente. L'organo principale è costituito da un'assemblea plenaria che si tiene 3-4 volte all'anno. La funzione di collegamento tra i componenti, e il punto di riferimento per gli esterni, è svolto dalla segreteria, che viene gestita a rotazione dalle istituzioni partecipanti, con un incarico annuale. Dal punto di vista operativo sono stati istituiti diversi gruppi di lavoro che si occupano di temi specifici e a cui partecipano i bibliotecari secondo le loro esigenze. Sono attivi cinque gruppi di lavoro: comunicazione³, formazione⁴, laboratori di lettura⁵, LOD, soggettario⁶.

Nel corso degli oltre dieci anni di attività il CoBiS ha dunque migliorato sensibilmente la fruibilità del patrimonio documentario conservato dalle biblioteche aderenti (oltre 1 milione e mezzo di volumi e più di 40.000 testate di periodici), orga-

- Al 31 agosto 2019 aderiscono al CoBiS: biblioteche afferenti ad Accademie, Biblioteche, Centri, Fondazioni, Istituti culturali e Musei. L'elenco e una descrizione sintetica di ciascuna biblioteca è disponibile sul sito web: https://cobis.to.it>.
- ³ Si occupa di migliorare la visibilità istituzionale del CoBiS e ha come obiettivo l'ampliamento della fruibilità delle risorse documentarie attraverso la promozione di iniziative e attività scientifiche delle biblioteche specialistiche aderenti (convegni, workshop, progetti specifici ecc.).
- ⁴ Si occupa di organizzare corsi di aggiornamento professionale specialistico con tematiche proposte dai bibliotecari del CoBiS (dal libro antico, alla conservazione delle collezioni, al diritto d'autore, alle REICAT, a Wikipedia e le biblioteche).
- Per promuovere la conoscenza sul territorio delle biblioteche specialistiche su proposta delle Biblioteche Civiche torinesi nel 2012 il CoBiS ha aderito all'iniziativa dei laboratori di lettura che si tengono generalmente in occasione della Giornata mondiale del Libro e delle biblioteche oppure in occasione di specifici eventi. I bibliotecari hanno quindi sperimentato la lettura di testi specialistici ad alunni della scuola primaria e secondaria.
- ⁶ Si confronta sui criteri di soggettazione dei documenti specialistici e dal 2010 ha una convenzione con la Biblioteca Nazionale Centrale di Firenze per l'invio di termini controllati.

nizzato momenti di aggiornamento professionale e creato una rete attiva – anche tra i bibliotecari. Tra gli esempi più significativi della capacità progettuale e di riflessione per i bibliotecari che aderiscono, sono il progetto CoBiS LOD e il progetto CoBiS Digital Library & Archives.

2. Il progetto CoBiS LOD

2.1 Le fasi iniziali

Alla fine del 2016 sei biblioteche del CoBiS, grazie a un contributo della Regione Piemonte – hanno partecipato al "progetto pilota". L'obiettivo del progetto (tuttora in corso) è la creazione di uno strumento, aperto ed efficiente per creare un catalogo collettivo "arricchito" da risorse esterne – anche internazionali – in maniera consequenziale e sempre in aggiornamento. L'esigenza è dovuta anche al fatto che le biblioteche aderenti usano almeno cinque gestionali diversi per la catalogazione del loro patrimonio librario (SeBiNa, SbnWeb, Clavis, Erasmonet e Bibliowin a seconda del polo di appartenenza oppure di scelte istituzionali)⁸ e Opac locali differenti.

La piattaforma *front-end*, uscita dalla fase pilota, è stata completamente ridefinita per farla evolvere da esito di un progetto-pilota a sito "in produzione" nel 2019. Oggi ospita i dati di 12 diverse biblioteche³ e permette al pubblico di interrogare un unico catalogo e allo stesso tempo di confrontarsi con informazioni provenienti da strumenti esterni nel web. Il passo successivo, che si sta intraprendendo nel momento in cui scriviamo, è quello di fare sinergia con i progetti di digitalizzazione condotti nel corso degli ultimi dieci anni in Piemonte e altrove. Il progetto LOD del CoBiS, un progetto di grande innovazione nel suo settore, è inoltre oggi il primo in Italia, e uno dei primi a livello internazionale, ad esporre un *endpoint* SPARQL secondo l'ontologia BIBFRAME 2.0, implementando nello stesso tempo gli standard di interoperabilità raccomandati da due degli enti più autorevoli in materia di standardizzazione, il W3C e la Library of Congress.

2.2 Perché i LOD?

Le tecnologie linked data rispondono alla crescente necessità di adottare strategie comuni di condivisione e valorizzazione dei patrimoni, e soprattutto di sfruttare

- Accademia delle Scienze di Torino, Associazione Archivio Storico Olivetti, Biblioteca Nazionale del Club Alpino Italiano, Deputazione Subalpina di storia patria, INAF Osservatorio Astrofisico di Torino, INRIM. L'Accademia delle Scienze che è un ente privato è stato individuato come capofila del progetto.
- 8 Alcune di esse non partecipano ad SBN e pertanto la ricerca bibliografica può essere assai difficoltosa e frammentata.
- ⁹ Alle seconda fase si sono aggiunte: Accademia di Medicina, Fondazione G. Amendola, Istituto per i beni musicali del Piemonte, Biblioteca civica musicale "A. Della Corte", Centro Internazionale di Studi Primo Levi - Polo del '900, Istituto di studi storici G. Salvemini.

l'architettura capillare del web per costituire una rete di conoscenza distribuita. I LOD sono una tecnologia nata proprio per favorire l'interoperabilità tecnica e semantica tra dati da fonti diverse.

La trasformazione in linked data rende i dati catalografici non solo interoperabili tra di loro, ma anche nei confronti di tutte le fonti esterne che utilizzano gli standard LOD raccomandati dal W3C¹⁰. Diverse istituzioni, come si vedrà, hanno iniziato ad *aprire* i propri cataloghi, a renderli accessibili in rete tramite standard di interoperabilità (come SPARQL).

Grazie all'interlinking tra le risorse del CoBiS e risorse analoghe afferenti a fonti LOD esterne, il portale del CoBiS è oggi in grado di esporre un *catalogo aumentato*, il cui contenuto informativo iniziale viene ampliato recuperando in tempo reale informazioni provenienti da database esterni, in continuo aggiornamento da parte della comunità web (come VIAF, Dbpedia, Wikidata e Internet Archive).

Il catalogo aumentato migliora nel tempo in maniera automatica, inoltre le biblioteche stesse sono incentivate a contribuire ai progetti di community (sia per mezzo di *editing* diretto o *upload* di risorse su Wikidata e/o Internet Archive, sia mediante l'istituzione manuale di collegamenti tra i propri cataloghi e tali piattaforme collaborative).

2.3 Alcuni aspetti tecnici del progetto¹¹

Punto di partenza dell'intero progetto è la trasformazione in LOD dei dati bibliografici provenienti dai 5 software gestionali in uso. I dati catalografici sono codificati non solo in diversi formati di file (XML, CSV e JSON) ma anche strutturati secondo formati standard non unici (MARC21, UNIMARC e Dublin Core OAl-PMH). La trasformazione in linked data richiede quindi che tutte le fonti dati siano rimodellate in RDF, lo strumento base proposto da W3C per la codifica, lo scambio e il riutilizzo di metadati strutturati che prevede che ogni informazione sia espressa sotto forma di triple soggetto-predicato-oggetto. Le informazioni in RDF vanno poi "modellate" secondo un' ontologia¹²: in ambito bibliografico, nel 2012 la Library of Congress ha pubblicato BIBFRAME, un modello di dati nato per favorire il passaggio dal modello catalografico basato su MARC al web semantico dei Linked Data. Con l'uscita della piattaforma dalla fase pilota, si è proceduto ad aggiornare l'ontologia da BIBFRAME 1.0 a BIBFRAME 2.0, che or-

¹⁰ Si veda in particolare: https://www.w3.org/DesignIssues/LinkedData.html.

Per gli aspetti più tecnici si veda Giorgio Bevilacqua, Interoperabilità e arricchimento semantico dei dati bibliografici: il progetto Linked Open Data del CoBiS, in: La biblioteca che cresce, contenuti e servizi tra frammentazione e integrazione, Milano, Editrice Bibliografica, 2019.

Sistema che stabilisce i criteri di realtà di determinati oggetti con il presupposto di un linguaggio formale.

ganizza le informazioni relative a una determinata risorsa sulla base di tre livelli (opera¹³, istanza¹⁴, item¹⁵)¹⁶.

La piattaforma attuale grazie a un processo di revisione conclusosi nel febbraio 2019, implementa una serie di procedure ottimizzate per essere facilmente replicabili, sia in termini di aggiornamenti periodici dei dati sia per facilitare l'acquisizione dei cataloghi di nuove biblioteche.

Per gli authority file, si procede a convertire in un unico grafo RDF le diverse fonti d'autorità, nell'ambito del quale è necessario poi riconciliare le entità duplicate, anche facendo affidamento alla presenza di codici identificativi esterni per la deduplicazione. Già nella fase pilota del progetto è stato avviato un processo di *interlinking* di identità tra l'authority file del CoBiS e soggetti analoghi presenti in fonti LOD esterne. Il processo impiega principalmente procedure automatiche che, basandosi sulla presenza di codici identificativi SBN su Wikidata, consentono allo stato attuale di interconnettere circa 4.000 autori del CoBiS con Wikidata stessa, VIAF e altre fonti esterne¹⁷.

2.4 Il Portale Web

Il Portale del progetto linked open data del CoBiS è pubblicato nel nuovo sito del CoBiS¹⁸ (Fig. 1).

La home page si presenta suddivisa in tre sezioni principali: una *ricerca generica* all'interno dei dati catalografici pubblicati in LOD, una ricerca "guidata" tra i *dati* catalografici principali (ovvero *Autori* e *Opere*), degli approfondimenti tematici o *percorsi*.

In tutte le pagine si trova poi una barra di navigazione che permette di accedere alle sottopagine del portale. In particolare nella sezione *Ontologia* (Fig. 2), è pos-

- ¹³ Il livello che riflette l'essenza concettuale della risorsa, legata a informazioni su titolo, autore e argomento.
- Permette di identificare una specifica manifestazione dell'opera, riportando ad esempio informazioni sull'editore, il luogo e la data di pubblicazione.
- ¹⁵ Questo livello (nuovo, rispetto alla versione precedente) identifica una specifica copia fisica o digitale di un'istanza, consentendo di registrare anche le informazioni relative alla collocazione della copia in una determinata biblioteca.
- Grazie al contributo della Regione Piemonte per la trasformazione dei dati dalle fonti originarie in BIBFRAME 2.0 è stato realizzato JARQL, un nuovo software open source che consente di prendere come input anche dati strutturati in forma di albero (come XML e JSON), i più frequenti tra le fonti dati del CoBiS.
- Per i casi in cui le procedure automatiche non danno risultati soddisfacenti (ancora numerosi) è stato sviluppato OLAF (Open Linked Authority File), un'interfaccia di crowd-sourcing semiautomatico che supporta gli esperti di dominio nell'attività di interlinking. OLAF suggerisce potenziali relazioni tra entità (nel nostro caso autori) simili. Le entità candidate sono sottoposte alla valutazione dei bibliotecari tramite un'interfaccia grafica, che consente loro di consultare le fonti web interconnesse. L'interazione tra l'automazione e la validazione manuale aumenta la qualità e affidabilità dei dati, abilitando i bibliotecari a contribuire direttamente alla linked open data cloud.
- ¹⁸ https://dati.cobis.to.it.



sibile fare delle *query* all'*endpoint* SPARKL e leggere una descrizione più tecnica del progetto.

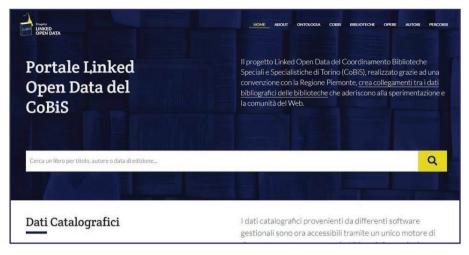


Figura 1. Home page del portale linked open data del CoBiS

Mentre nella sezione *Biblioteche* si può navigare tra le schede che illustrano le biblioteche partecipanti al progetto e i loro patrimoni bibliografici.

In questa sezione agli utenti particolarmente abili, il catalogo aumentato consente di estendere semanticamente il contenuto informativo dell'authority file, acquisendo notizie sul contesto biografico degli autori. Se opportunamente strutturate tali informazioni aggiuntive possono fornire nuove modalità di accesso al ca-

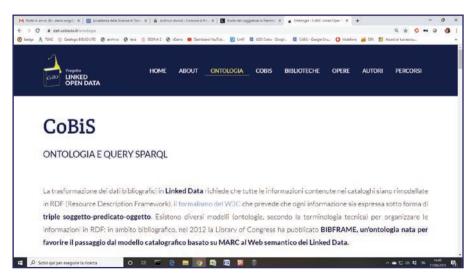


Figura 2. Ontologia

talogo. È possibile ad esempio, attraverso l'implementazione di query SPARQL federate¹⁹, realizzare una mappa dei luoghi in cui hanno studiato e lavorato gli autori presenti (reperendo da Wikidata l'informazione su quali Università e Accademie abbiano frequentato gli autori e le corrispondenti coordinate geografiche). La mappa fornisce, una visualizzazione interessante di quali sono stati i centri di formazione più frequentati e può essere utilizzata come strumento alternativo di accesso al catalogo.

Le sottopagine *Autori* e *Opere* (Fig. 3) sono particolarmente utili per l'utenza poiché permettono di applicare dei filtri alla ricerca che si intende fare. Anche queste due sottopagine sono suddivise in tre sezioni. Nella prima sezione si trova un rettangolo di ricerca specifico che restringe la ricerca dell'utente al solo campo autore o al solo campo del titolo dell'opera. Ciò permette ad esempio di cercare più agevolmente le opere *di* Alessandro Manzoni.

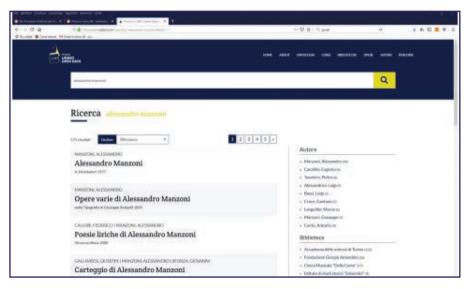


Figura 3. Sottopagine autore-opere

Oppure quelle *su* Alessandro Manzoni (o per meglio dire nel cui titolo si trovano le parole Alessandro *e* Manzoni).

La sottopagina *Autori* risulta poi suddivisa in due sezioni: una stringa di ricerca semplice e una sezione *Autori* in evidenza (fig. 4) con un box grafico generato dinamicamente recuperando informazioni distribuite nel web tramite l'esecuzione di query SPARQL²⁰. In questo momento si nota dal grafo che gli autori legati alla mu-

¹⁹ Ricerche che combinano vari quesiti insieme (Es. quanti Autori nati in Italia hanno studiato in Europa ecc.).

²⁰ Anche la pagina delle opere viene generata automaticamente tramite l'esecuzione di query SPARQL.

sica hanno la predominanza iconografica tra gli autori in evidenza. Ciò è dovuto non alla quantità bensì alla pluralità di progetti linked data relativi alla musica che ha permesso di rendere questi personaggi più ricchi di informazioni presenti sul web rispetto ad autori scientifici o letterari.

Muovendosi col cursore su una delle immagini è possibile visualizzare nome dell'autore, ritratto, estremi cronologici e numero di sue opere²¹ presenti nel portale.



Figura 4. Autori in evidenza

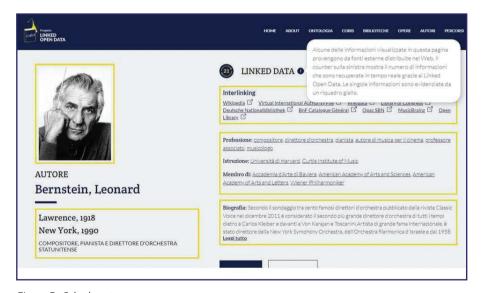


Figura 5. Scheda autore

²¹ Il numero non indica il quantitativo di edizioni ma le copie fisiche.

La "scheda" dell'autore (Fig. 5) risulta particolarmente dettagliata e ricca di dati provenienti da fonti esterne.

Tutto ciò che è riquadrato in giallo non è presente sul catalogo di origine ma è risultato del catalogo aumentato.

In particolare, il nome e l'elenco dei libri provengono dall'endpoint del CoBiS, la biografia da SBN o Dbpedia, l'immagine da WikiCommons, i metadati biografici da Wikidata; inoltre sono elencate nell'interlinking una serie di risorse interconnesse

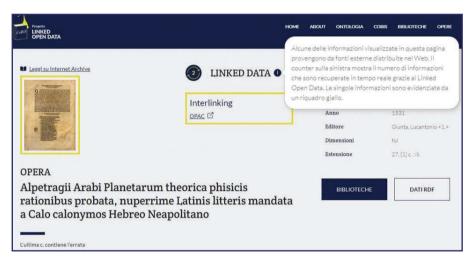


Figura 6. Scheda opere

che rappresentano altre potenziali fonti di dati (VIAF, Library of Congress, Deutsche National Bibliothek, Bibliothèque Nationale de France, Servizio Bibliotecario Nazionale, Dizionario Biografico degli Italiani Treccani).

La "scheda" delle opere (Fig. 6) si presenta invece suddivisa in tre parti: un'immagine del frontespizio se disponibile su web, il titolo e l'autore; dati recuperati dal data cloud.

In alcuni casi si è in grado di consultare direttamente su Internet Archive una copia digitale dell'opera, sfruttando così appieno le potenzialità degli open data.

Cliccando sul bottone RDF, tutte le triple che descrivono la risorsa possono essere visionate direttamente: le pagine web sono generate dall'esecuzione in tempo reale di query SPARQL.

2.5 Dati aperti, attività di community e sviluppi futuri

L'architettura linked data *full stack* del progetto fa sì che al miglioramento e alla crescita dei dati di community corrisponda un'evoluzione del portale. Per questo motivo nelle prossime fasi del progetto, che avranno luogo tra la fine del 2019 e i primi mesi del 2020, si intende ottimizzare e intensificare le procedure di interlinking automatico ed editing verso piattaforme collaborative, soprattutto Wikidata e

Internet Archive, nonché estendere le funzionalità del front-end, in modo da rendere più esplicita l'architettura LOD e le sue possibilità di ampliamento del contenuto informativo del catalogo.

Tra gli sviluppi futuri del progetto c'è, come si vedrà qui di seguito, l'avviamento di una convergenza tra il progetto LOD e un progetto di digitalizzazione in corso da parte del CoBiS.

3. Il progetto "CoBiS Digital Library & Archives"

Il progetto "CoBiS Digital Library & Archives" (digitalizzazione di fondi bibliografici e archivistici appartenenti alle raccolte di istituti afferenti al CoBiS) è stato finanziato con specifico contributo della Regione Piemonte; avviato nell'aprile 2018 con la stesura del capitolato e la pubblicazione della richiesta di offerta sul sito del MEPA, ha riguardato 17 istituti²².

Precipue finalità del progetto sono state: garantire la tutela e la conservazione dei documenti oggetto di intervento e favorire la loro consultazione e diffusione attraverso la digitalizzazione, il riconoscimento ottico dei caratteri e la pubblicazione dei *file* sia sulle piattaforme digitali dei singoli istituti sia sulla Teca Digitale Piemontese (o sua evoluzione).

L'obiettivo stabilito in fase iniziale di pubblicare online il materiale documentario avrà come naturale conseguenza la possibilità di mettere in contatto i ricercatori che consulteranno tali risorse con le istituzioni (talvolta poco conosciute) che tali documenti conservano.

Il ruolo di ente capofila nel coordinamento del progetto è stato affidato alla Biblioteca civica Centrale di Torino in virtù della consistenza del materiale documentario presente nelle sue raccolte, della diffusione capillare del Sistema bibliotecario urbano (cui la Biblioteca afferisce) sul territorio cittadino e della consequente possibilità di raggiungere un pubblico vasto e differenziato.

3.1 La selezione dei documenti da digitalizzare

Il materiale da sottoporre a digitalizzazione è stato selezionato da parte delle diverse biblioteche in ragione di diverse istanze:

Accademia Albertina di Belle Arti, Accademia delle Scienze, Accademia di Agricoltura, Associazione Archivio storico Olivetti, Bibliomediateca Mario Gromo, Biblioteca civica Centrale di Torino, Biblioteca civica musicale "Andrea Della Corte" di Torino, Biblioteca nazionale del Club Alpino Italiano, Centro Studi Piemontesi, Centro Studi Sereno Regis, Deputazione Subalpina di storia patria, Educatorio della Provvidenza, Fondazione Giorgio Amendola, Fondazione Tancredi di Barolo, Museo del Risorgimento, Soprintendenza Archeologia Belle Arti e Paesaggio per la Città Metropolitana di Torino, SPABA - Società Piemontese di Archeologia e Belle Arti. Inizialmente l'elenco comprendeva anche l'Accademia di Medicina, il Centro Studi Vietnamiti e l'Istituto dei Sordi, che hanno abbandonato il progetto prima del passaggio alla fase attuativa.

- rarità e specificità della documentazione, caratteristiche che impediscono il reperimento di esemplari analoghi all'interno di altre biblioteche o di altre piattaforme. Tale unicità è ancor più evidente nel caso di materiale archivistico o, comunque, manoscritto;
- rarità e specificità del materiale nell'ambito della complessiva raccolta documentaria di un singolo istituto, utile a comprenderne e giustificarne l'origine e l'importanza;
- difficoltà di consultazione dovuta a molteplici ragioni: dalla fragilità dei supporti alla difficoltà di accesso fisico ai locali;
- importanza del materiale all'interno del contesto locale (storia del territorio, della società, delle tecnologie, delle istituzioni);
- presenza di materiale analogo proveniente dal medesimo fondo e già digitalizzato nell'ambito di precedenti progetti. In tal caso l'intervento si rende necessario al fine di completare lavori rimasti in sospeso.

Dal punto di vista catalografico, è stato richiesto che tutte le opere a stampa selezionate fossero catalogate e presenti in SBN.

3.2 La richiesta di offerta pubblicata sul Mercato elettronico della pubblica amministrazione

La richiesta di offerta è stata formulata in base a sette punti, corrispondenti alle diverse fasi di esecuzione del lavoro:

- digitalizzazione e produzione dei file immagine nei formati TIFF e JPEG;
- controllo delle immagini scansionate al fine di verificarne qualità e completezza;
- attribuzione della nomenclatura univoca delle immagini secondo gli standard previsti dalla Teca Digitale Piemontese;
- riconoscimento ottico del contenuto testuale (OCR) del materiale a stampa scansionato;
- creazione dei file PDF e dei file testuali del materiale librario e archivistico scansionato;
- produzione dei metadati in formato xml Maq 2.0.1 e Mets;
- riversamento delle immagini digitali e dei corrispettivi metadati su supporti di memorizzazione.

L'aggiudicazione è stata effettuata sulla base del criterio dell'offerta economicamente più vantaggiosa con il metodo aggregativo compensatore. Il punteggio totale massimo di 100 punti è stato ripartito in 30 punti per l'offerta economica e in 70 punti per l'offerta tecnica. I criteri di valutazione hanno individuato punteggi diversi da attribuire a ciascuna delle seguenti voci:

 modello organizzativo per la gestione del servizio, con particolare riferimento agli aspetti innovativi del processo di produzione (max punti 25);

- esperienze pregresse (max punti 5);
- infrastruttura tecnologica adottata (max punti 15);
- tempi di avvio, di realizzazione e di collaudo (max punti 10);
- proposte migliorative dei livelli di servizio proposti in capitolato (max punti 15).

I lavori sono stati svolti presso due diverse sedi: il Centro Rete del Sistema bibliotecario urbano della Città di Torino e la Biblioteca dell'Accademia Albertina di Belle Arti.

Per ogni documento sono state richieste immagini nei formati: TIFF 6.0 non compresso con risoluzione di 600 dpi ottici (non interpolati) e una profondità di colore di 24 bit RGB; JPEG sottocampionato a 300 dpi compresso e una profondità di colore di 24 bit RGB; JPEG sottocampionato a 150 dpi compresso e una profondità di colore di 24 bit RGB.

I documenti a stampa digitalizzati in formato TIFF sono stati sottoposti a riconoscimento ottico dei caratteri. I file di testo così prodotti sono stati adoperati per la generazione di file PDF multipagine, al cui interno è possibile sia la funzione di ricerca testuale sia la messa in evidenza delle parole ricercate.

I metadati sono stati realizzati secondo le linee guida della Regione Piemonte e dai metadati Mag è stata ottenuta la conversione in metadati Mets.

Per quanto concerne i macchinari, sono stati utilizzati più scanner a planetario professionali, con un campo di ripresa dal formato A5 al formato A2, in conformità a quanto previsto dalla normativa sulla tutela del materiale cartaceo pregiato.

Al momento della consegna la ditta aggiudicataria fornirà a ciascuna singola Biblioteca aderente al progetto un hard disk esterno USB-3, su cui avrà riversato immagini e dati relativi al materiale di sua pertinenza. I dati gestiti nell'ambito del progetto rimarranno di proprietà esclusiva dei singoli enti partecipanti, che avranno la possibilità di inserire le proprie scansioni anche sulle piattaforme interne all'istituto.

3.3 Il lavoro svolto

Al momento attuale i lavori di scansione sono stati ultimati, ma il progetto non è ancora fruibile in tutte le sue parti. È già possibile, tuttavia, formulare alcune considerazioni di massima relative alle difficoltà incontrate e ai risultati raggiunti. Dal punto di vista metodologico è stata importantissima la fase di preparazione del lavoro che, sebbene abbia richiesto diverso tempo, ha permesso di evitare interruzioni e ripensamenti nel corso delle fasi successive. A ciascun istituto sono stati richiesti fin dall'inizio gli elenchi del materiale, in modo da controllare sulle varie piattaforme che non esistessero medesime edizioni di una stessa opera già scansionate nel corso di altri progetti (in tal caso le opere già presenti in elenco potevano essere sostituite o sottoposte a nuova lavorazione qualora si fosse trattato di esemplari con note a margine, postille, tavole acquerellate, ecc.). Ciascuna

notizia bibliografica è stata controllata e confrontata con l'esemplare in possesso della biblioteca proponente e alle liste di volumi proposti è stata aggiunta l'indicazione del BID SBN. Sicuramente, una lunga esperienza di catalogazione da parte di alcuni bibliotecari degli enti partecipanti ha costituito un validissimo supporto allo svolgimento dei lavori e si è rivelato un valore aggiunto anche in virtù della positiva relazione di fiducia e collaborazione instaurata con il personale della ditta appaltatrice.

La possibilità di disporre di una visione d'insieme, offerta dagli elenchi preparatori in una fase di molto antecedente l'inizio dei lavori, ha posto gradualmente in luce alcune peculiarità presenti nelle raccolte dei singoli istituti, caratteristiche che in un secondo tempo si è deciso di privilegiare nel trattamento.

Infatti, in ragione del fatto che era stato presentato un progetto comune, era necessario convincere la Regione che il risultato sarebbe stato non soltanto sod-disfacente dal punto di vista tecnico e di salvaguardia dei documenti, ma anche sotto il profilo della valorizzazione del territorio. Una valorizzazione che in questo caso si declinava in una visibilità del tutto nuova di cui sarebbero stati oggetto enti assolutamente di valore, necessari nel determinare la ricchezza culturale del Piemonte ma poco conosciuti al di fuori della cerchia specialistica di riferimento.

E così, pur nella totale libertà concessa a ciascun istituto di proporre autonomamente su quali documenti intervenire, via via sono emerse proposte volte a privilegiare il materiale che ciascun ente ha prodotto nel corso della sua esistenza e che ha contribuito ad accrescerne il prestigio: gli annuari e i periodici di istituto, le pubblicazioni più importanti (a questo proposito è stato possibile superare il divieto di digitalizzare e rendere pubblico materiale ancora sotto diritti perché le liberatorie sono pervenute proprio dall'ente proprietario dei diritti medesimi).

In tal modo, a breve sarà possibile consultare da remoto *Studi piemontesi*, il *Bollettino della Società Piemontese di Archeologia e Belle arti*, il *Bollettino storico bibliografico subalpino*, gli annuari dell'Accademia di Agricoltura, i giornali di fabbrica Olivetti, i notiziari del Club Alpino Italiano, i periodici *Il Rinnovamento* (edito dalla Fondazione Giorgio Amendola) e *Il Risorgimento* (di proprietà della Biblioteca del Museo del Risorgimento).

Sarà parimenti disponibile la raccolta di disegni originali dello scenografo Giacomo Pregliasco di proprietà della Biblioteca civica Centrale di Torino, il fondo di bozzetti teatrali Ezio Cellini della Biblioteca civica musicale "Andrea Della Corte" di Torino e un fondo archivistico, il fondo Vico, che riunisce manoscritti e materiali ottocenteschi sulla Galleria Sabauda (di proprietà della Soprintendenza Archeologia Belle Arti e Paesaggio della Città Metropolitana di Torino). E ancora, ai primi periodici pubblicati dal movimento nonviolento in Italia (le cui liberatorie sono state chieste all'ente proprietario dei diritti direttamente dal Centro Sereno Regis) andranno ad affiancarsi gli abbecedari ottocenteschi del

Museo della Scuola, alcune raccolte di periodici di proprietà dell'Accademia delle Scienze e dell'Educatorio della Provvidenza e preziose monografie a carattere scientifico e didattico conservate dalla Bibliomediateca Gromo e dall'Accademia Albertina di Belle Arti.

Con la locale Soprintendenza bibliografica ed archivistica si è instaurata una proficua collaborazione; la funzionaria incaricata ha seguito tutte le fasi, dalla richiesta di permesso (accompagnata da sopralluoghi presso ciascuna sede) alla movimentazione del materiale.

La stipula della polizza assicurativa – necessaria nell'ambito di un progetto in cui il materiale da spostare era ingente e riconducibile a 17 istituti, è stata sottoscritta dalla ditta vincitrice del bando, che si è presa cura in toto del trasporto e della salvaguardia del materiale in fase di lavorazione (la messa in sicurezza dei locali è stata a carico della stazione appaltante).

3.4 L'integrazione con il progetto LOD

Il progetto CoBiS Digital Library & Archives ha previsto la digitalizzazione di una serie di opere appartenenti a sei biblioteche (sempre afferenti al CoBiS) aderenti anche al progetto LOD²³.

La proposta di avviare una convergenza fra i due progetti è emersa a seguito di ripetuti incontri con i funzionari della Regione Piemonte. In questa sede è stato concordato di riversare in prima istanza i documenti digitalizzati, anziché sulla Teca Digitale Piemontese o sua evoluzione, sulla piattaforma Internet Archive²⁴, già utilizzata come *interlinking* automatico dal progetto LOD.

A tale scopo saranno necessari i seguenti interventi:

- creazione di una collezione "CoBiS Digital Library & Archives" su Internet Archive;
- caricamento dei contenuti da parte di un utente facilmente riconducibile al progetto;
- definizione di una modalità condivisa per annotare la corrispondenza tra i BID SBN e gli URL di Internet Archive.

Nel corso dei prossimi mesi sarà pertanto possibile arricchire il portale CoBiS LOD anche di immagini digitalizzate provenienti da biblioteche interne al coordinamento.

²³ Accademia delle Scienze, Associazione Archivio storico Olivetti, Biblioteca civica musicale "Andrea Della Corte" di Torino, Biblioteca nazionale del Club Alpino Italiano, Deputazione Subalpina di storia patria, Fondazione Giorgio Amendola.

²⁴ Il riversamento sulla Teca Digitale Piemontese verrà comunque effettuato successivamente.

4. I linked open data all'estero e in Italia

4.1 L'esperienza internazionale

Il gruppo di lavoro sulla comunicazione del CoBiS durante il convegno del 2014 ("Le biblioteche specialistiche un valore tra realtà e futuro")²⁵ ha intrapreso lo studio sui linked open data analizzando le esperienze italiane e internazionali. Ancora oggi il progetto di pubblicazione e comunicazione dei dati della Bibliothèque nationale de France (BnF) rimane un esempio singolare. Il progetto nasce nel 2009 dall'esigenza di rendere visibili contenuti digitali di Gallica. Ora data.bnf.fr (Fig. 7) è un'interfaccia web in linked open data, che comprende il catalogo generale, il catalogo degli archivi e manoscritti e Gallica. A questi si aggiungo i legati e i dati provenienti dal catalogo collettivo delle biblioteche di ricerca francesi, WorldCat, Viaf e Wikipedia. In questo modo l'utente ha accesso sia al dato catalografico, ad altre risorse esterne e infine, se esiste, al documento originale.

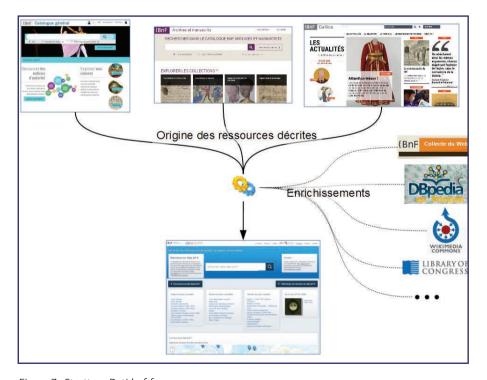


Figura 7. Struttura Dati.bnf.fr

²⁵ https://cobis.to.it/il-catalogo-aumentato-come-cambia-oggi-il-modo-di-accedere-ai-documenti-e-alle-informazioni/>.

Secondo i dati statistici²⁶ è interessante rilevare che durante il suo primo anno di funzionamento, *data.bnf.fr* ha raggiunto 637.650 visitatori individuali e 1,2 milioni di visite alle pagine. È significativo rilevare che l'80,6% delle visite provengono da un motore di ricerca web. Questo conferma che la maggior parte degli utenti del web utilizza motori di ricerca generali per qualunque tipo di bisogno, così come è ormai assodato che la maggioranza degli utenti del web utilizza, in prima battuta, *wikipedia* per soddisfare i propri bisogni informativi. Il progetto francese coopera con altre biblioteche nazionali impegnate nei LOD come Deutsche Nationalbibliothek, British Library e Library of Congress o con VIAF.

Parimenti la Cornell University Library, Harvard Library Innovation Lab e Stanford University Libraries, continuano la sperimentazione, sostenute da cospicui finanziamenti sul progetto Linked Data for Libraries Project (LD4L), l'obiettivo è quello di creare un sistema che estrae le informazioni delle singole biblioteche accademiche attraverso i LOD e che unifichi i dati². Il grande contenitore europeo Europeana è invece un progetto che utilizza i LOD per valorizzare il patrimonio culturale europeo. Alla base vi è l'idea di raccogliere le collezioni di biblioteche, musei e archivi, attraverso uno schema comune di metadati, e utilizzare un'unica piattaforma di indicizzazione e accesso ai contenuti. Il portale è il servizio chiave della Fondazione Europeana, che ha come obiettivo l'aggregazione e la pubblicazione dei dati creati dalle istituzioni culturali e dalle comunità scientifiche.

4.2 L'esperienza italiana

Il CoBiS studia i linked open data (LOD) sia a livello internazionale che italiano, orientandosi in particolare sui set di dati che includono i metadati dei libri, monografie e periodici²⁸.

In Italia si sono aperte piattaforme di dati pubblicati come dati aperti a seguito anche del recepimento da parte dell'Italia delle direttive comunitarie per le pubbliche amministrazioni per opera dell'art. 52 del d.lgs 7.3.2005 n. 82²⁹, poi con l'art. 9 del d.l n. 179/2012³⁰.

²⁶ Valeria Lo Castro, *Linked data nelle biblioteche digitali e di ricerca*: una panoramica internazionale, «Biblioteche Oggi», 33 (2015), gennaio – febbraio, p. 36-44,

< http://www.bibliotecheoggi.it/rivista/article/view/5>.

²⁷ Ibidem.

²⁸ Gabriella Morabito - Anna Maria Viotto, Giornata di studio delle biblioteche specialistiche di Torino e area metropolitana, «AIB notizie», 30 novembre 2014, < http://aibnotizie.aib.it/piemontecobis/>.

²⁹ https://www.agid.gov.it/sites/default/files/repository_files/leggi_decreti_direttive/dl-7-marzo-2005-82 0.pdf.

https://www.gazzettaufficiale.it/moduli/DL_181012_179.pdf.

Le esperienze italiane (Fig. 8)³¹ fanno riferimento a grandi istituzioni, ad esempio, l'ISTAT³², l'ISPRA³³, il CNR³⁴, la Camera dei Deputati³⁵, il Senato della Repubblica³⁶, iniziative che hanno come obbiettivo di fornire accesso pubblico alle loro informazioni.



Figura 8. Progetti in Italia sui LOD

I LOD per i beni culturali rappresentano un'importante fonte per il sapere e per la partecipazione collettiva alla costruzione di nuovi contenuti culturali.

Questo processo di associazione dei dati sui patrimoni archivistici, bibliotecari e museali, e il loro collegamento anche a fonti esterne, ha visto nascere molti progetti e casi studio, come Culturaltalia³⁷, l'aggregatore nazionale del patrimonio digitale. Un altro gruppo di lavoro è quello dell'Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU) che sta portando avanti un'attività di ricerca per la conversione dei record di SBN in linked open data e che ha recentemente presentato il suo progetto che comprende il rilascio dei dati bibliografici in LOD³⁸.

Un secondo progetto è quello degli Istituti campani che si rivolge sia al patrimonio archivistico che bibliotecario digitalizzato³⁹. Per il CoBiS fonte di ispirazione è sta-

³¹ Silvia Mazzini, Regesta.exe.

³² http://datiopen.istat.it/.

³³ http://dati.isprambiente.it/.

³⁴ http://data.cnr.it/site/.

³⁵ http://dati.camera.it/it/linked-data/.

³⁶ http://dati.senato.it/sito/home.

³⁷ http://www.culturaitalia.it/.

https://www.iccu.sbn.it/export/sites/iccu/documenti/2018/SRI_CG_DLG_CTecnico_ICCU_2018.

³⁹ http://www.polodigitalenapoli.it/.

to, anche, il progetto degli archivi digitali della Shoah del CDEC⁴⁰, che permette di far incrociare le banche dati dei nomi delle vittime.

Altri casi studio sono Share Catalogue⁴¹, una piattaforma con cataloghi bibliografici che segue il data model BIBFRAME Linked Data, un punto di accesso al patrimonio di beni integrati (libri, documenti, ecc.) e il Museo Galileo di Firenze⁴², una banca dati cumulativa che consente la consultazione integrata delle collezioni, delle bibliografie e dei data base prodotti nell'ambito dei diversi progetti di studio e ricerca. Comprende attualmente circa 330.000 record relativi ai testi a stampa posseduti dalla biblioteca.

Un recente progetto, del 2018, è quello dell'Archivio del Quirinale che pubblica dati liberi e riusabili in LOD, dell'attività della Presidenza della Repubblica e Segretario generale dal 1948⁴³.

5. Conclusioni

Sicuramente il progetto CoBiS LOD è stato e continua ad essere una sfida per le biblioteche partecipanti. All'inizio è nato come un momento di formazione per i bibliotecari convinti che i linked open data possano offrire nuove risorse informative, cambiando la percezione delle biblioteche specialistiche e modificando in parte il ruolo dei bibliotecari stessi («Le biblioteche stanno portando sul web una prospettiva a lungo termine. Stanno raccogliendo libri da secoli e dati da decenni»)44. È indubbio che intraprendere questa strada ha consentito di far conoscere o riscoprire le biblioteche specialistiche, spesso tacciate di essere luoghi chiusi e poco inclini alle innovazioni. Tuttavia persistono notevoli difficoltà, di varia natura, legate al progetto in corso. Reperire la dotazione finanziaria per l'avvio della fase pilota non è stato semplice e la sensibilità dimostrata dalla Regione Piemonte è giunta dopo numerevoli e compresivi incontri dove sulla carta veniva illustrato un progetto futuristico. Il problema economico permane soprattutto in merito al proseguo del progetto, ma anche sulla gestione e il mantenimento di daticobis. I progetti che abbiamo studiato e a cui ci siamo ispirati sono tutti supportati da ingenti risorse finanziare e tecniche. Dal lato tecnico siamo consapevoli di alcune criticità, prima fra tutte la scarsa indicizzazione del sito daticobis da parte dei motori di ricerca (Google), siamo al momento poco visibili verso l'esterno e in attesa che l'ICCU rilasci i dati Sbn in formato LOD pensiamo di forzare il

 $^{^{\}mbox{\tiny 40}}$ http://www.cdec.it/home2_2.asp?idtesto=940&idtesto1=1488&lemma=lod.

http://catalogo.share-cat.unina.it/sharecat/clusters.

⁴² https://www.museogalileo.it/it/biblioteca-e-istituto-di-ricerca/biblioteca-e-archivi/banca-daticumulativa.html.

⁴³ https://archivio.guirinale.it/aspr/redazione/linked-open-data.

⁴⁴ Romain Wenz, *Linked open data per i nuovi servizi bibliotecari: l'esempio di data.bnf.fr*, «JLIS.it», 4 (2013), n. 1, p. 403-415, https://www.jlis.it/article/view/5509>.

collegamento a pagine web dall'esterno con la creazioni di voci su wikidata e wi-kipedia. La convergenza con il progetto CoBiS Digital Library dovrebbe supportare notevolmente questo processo.

Si aprono poi considerazioni sul ruolo della professione di bibliotecario.

La comunità dei LOD coinvolge gruppi di lavoro con diverse competenze, bibliote-cari, ricercatori, scienziati, informatici, filosofi. I progetti linked open data favori-scono l'integrazione e la produzione di dati catalografici non prodotti dai bibliote-cari, è probabile che a questi ultimi si chieda competenza sull'authority control «la comunità professionale bibliotecaria dovrà svolgere un ruolo fondamentale perché la sola che possiede oggi le competenze necessarie per garantire integrità e univocità dei dati»⁴⁵.

The LOD (linked open data) experience of the CoBiS currently includes 12 libraries that are part of the Coordination of special and specialized libraries in the Turin metropolitan area. The project, financed by the Piedmont Region, provides for the sharing of the cataloging data of the individual libraries on the portal https://dati.cobis.to.it/. A new online site has thus been created which thanks to linked open data technology allows the user to search for catalog data but also to acquire new information, thanks to links to other external sources such as, for example, wikipedia. The digitization of bibliographic and archival collections of institutes related to the CoBiS (CoBiS Digital Library & Archives) will soon be integrated into the CoBiS LOD project: an improvement of the service for users in the Italian and foreign panorama of experimentation related to linked open data.

L'ultima consultazione dei siti web è avvenuta nel mese di dicembre 2019

⁴⁵ Carlo Bianchi, Library linked data e il futuro delle biblioteche, 6 marzo 2012, https://sites.google.com/site/homepagecarlobianchini/Ricerca/library-linked-data-e-il-futuro-delle-biblioteche.