

OSSERVATORIO

L'universo delle risorse culturali: lampi di genio e azioni concrete

Lightning talks presentati al Convegno AIB CILW 2016

a cura del Gruppo di studio AIB
Catalogazione, indicizzazione, linked open data e web semantico (CILW)

Premessa

di Andrea Marchitelli

Si è tenuto a Roma, lo scorso 21 ottobre 2016, presso la sala conferenze della Biblioteca nazionale centrale “Vittorio Emanuele II” di Roma, il Convegno AIB CILW 2016, “La rinascita delle risorse dell’informazione: granularità, interoperabilità e integrazione dei dati”, organizzato dal Gruppo di studio AIB Catalogazione, indicizzazione, linked open data e web semantico (CILW), dall’Associazione italiana biblioteche, Sezione Lazio e dalla Biblioteca nazionale centrale di Roma stessa.

Il Convegno AIB CILW 2016 mirava, innanzitutto, a realizzare un momento di incontro tra esperienze, ricerche e studi in corso sui temi dell’innovazione teorica, metodologica, tecnologica e professionale in area MAB e LAMMS, con particolare riguardo alle tecnologie dei LOD e al progetto del Web semantico.

La prima parte del convegno, impostata in maniera tradizionale, era volta a proporre una riflessione teorica, organica e innovativa, in grado di chiarire i principali aspetti del tema della convergenza delle risorse informative tramite i dati collegati.

La seconda parte – che ho avuto l’onore e il piacere di coordinare – ha visto la raccolta di contributi attraverso una *call for proposals* aperta, tra i quali sono stati selezionati i sette presentati in forma di *lightning talk*, sei dei quali vedono ora luce, in modalità più tradi-

GRUPPO DI STUDIO AIB CILW (Catalogazione, indicizzazione, linked open data e web semantico), composto da Roberto Raieli, Marinella Cisternino, Elena Corradini, Antonella Iacono, Antonella Trombone, sito web <www.aib.it/struttura/commissioni-e-gruppi/cnc>, e-mail cnc@aib.it.

Home page del Convegno AIB CILW 2016: <<http://www.aib.it/attivita/congressi/c2016/giornata-studi-aib-cilw-2016/>>.

ANDREA MARCHITELLI, EBSCO information services, e-mail amarchitelli@ebSCO.com.



zionale, nelle pagine che seguono. Questa sessione puntava a dare l'opportunità ad altri studiosi di esporre le proprie idee ed esperienze sui temi trattati, in particolare illustrando i risultati di progetti, conclusi o ancora in corso, legati alle tematiche congressuali.

La terza parte, infine, era orientata alla discussione di ulteriori progetti ed esperienze in grado di esemplificare le potenzialità applicative dei LOD e del Web dei dati.

Il 'lightning talk' è una modalità espositiva per interventi presentati a congresso piuttosto nuova, almeno in Italia per il nostro settore, ma con caratteristiche che la rendono interessante e potenzialmente assai efficace. Si tratta, infatti, di brevi relazioni, della durata di pochi minuti: in genere 5, come è stato anche per il convegno CILW. In questa manciata di minuti, pochissimi rispetto ai 20-30 di una presentazione tradizionale, il relatore ha l'opportunità, e il compito, di veicolare i suoi contenuti nella maniera più chiara possibile.

I relatori della sessione *lightning talks* del convegno CILW hanno tutti interpretato correttamente l'idea fornendo, ciascuno nel tempo limite di 5 minuti, presentazioni molto convincenti, espresse in maniera semplice e dinamica, e corredati di un buon apparato di immagini e slide (quando non anche video, in due casi su sette)¹. Tolti i fronzoli a causa del tempo ridotto, ciascun contributo è emerso per gli aspetti centrali del suo contenuto, pochi, ma tutti messi nel giusto rilievo e ben veicolati alla platea.

I primi due interventi presentati (*Criteria for the selection of a semantic repository for managing SKOS data* e *Modelli per la comunicazione dei dati di ricerca in archeologia: il caso dei Getty vocabularies come linked open data*) hanno posto il focus su uno degli aspetti centrali del progetto Web semantico e dei LOD, quello dell'interoperabilità semantica, gestibile attraverso SKOS, di cui si è mostrata un'applicazione creata per gestire in maniera semplice query di questo tipo, e una realizzazione concreta, quella della distribuzione dei *Getty vocabularies as linked open data* attraverso LOD.

I successivi due interventi (*Coordinamento delle biblioteche speciali e specialistiche di Torino: progetto linked open data* e *Gli authority file per l'integrazione cross-domain dei beni culturali: riflessioni su un approccio alla lettura trasversale dei beni culturali della Chiesa cattolica italiana*) hanno mostrato le potenzialità di ontologie e LOD riguardo all'integrazione e interoperabilità di dati provenienti da strumenti e fonti differenti, catalogati e descritti in modalità differenti e riguardanti oggetti differenti (beni librari, archivistici, opere d'arte, persone, enti...).

I successivi due interventi hanno mostrato due progetti in corso di realizzazione, che hanno entrambi l'obiettivo di mettere a disposizione del pubblico, attraverso i LOD, una specifica raccolta (*L'archivio fotografico Zeri in linked open data*) o un intero museo (*U. Porto digital museum: towards convergence in university's heritage resources management*)².

L'ultimo intervento (*Coming AUTH: per una bonifica e implementazione dell'authority file di SBN*) ha proposto un progetto volto all'implementazione dell'*authority file* di SBN, con una modalità professionale e una modalità aperta al mondo non bibliotecario, che avverrebbe tramite un servizio *Mix'n'match* come quello sviluppato da Wikimedia, invitando gli utenti, previa registrazione, a confermare la corrispondenza o meno di entità di diversa natura tratte da fonti diverse.

Insomma, una ricchezza di contenuti ed esperienza, che i relatori hanno efficacemente condiviso con il pubblico presente... nello spazio di poco più di mezz'ora!

¹ I video e le slide utilizzati durante la sessione dei *lightning talks* sono accessibili a partire dalla pagina web del programma del Convegno: <<http://www.aib.it/attivita/congressi/aib-cilw-2016-conferenza/2016/56896-convegno-aib-cilw-2016-programma>>.

² Questo *talk* non è tra quelli qui pubblicati.

Criteria per la selezione di un archivio semantico per la gestione di dati SKOS

di Ricardo Eito-Brun

Introduzione

Biblioteche, archivi e altre organizzazioni dedicate alla conservazione e alla disseminazione del patrimonio culturale hanno cominciato a porre attenzione alle tecnologie del Web semantico. Tali tecnologie fanno parte delle strategie per l'introduzione dei linked open data (LOD) e sono diventate una componente chiave dell'infrastruttura per la gestione delle informazioni di queste istituzioni.

Le tecnologie e gli standard del Web semantico includono, fra gli altri, i linguaggi RDF (Resource description framework) e RDF(S), OWL, RDFa e SKOS (Simple knowledge organization system)¹. Quest'ultimo è lo standard W₃C per codificare i vari tipi di sistemi di organizzazione della conoscenza: le classificazioni, i tesauri, le liste di intestazione per soggetto eccetera. Utilizzando lo standard SKOS, le istituzioni possono codificare e taggare gli strumenti di rappresentazione della conoscenza e utilizzarli nel Web per dare più facilmente accesso alle informazioni. Lo SKOS ha favorito l'opportunità di riutilizzare vocabolari e tesauri esistenti nel contesto delle nuove applicazioni LOD. Grazie a questo standard, si valorizza l'impegno profuso in passato per ideare e sviluppare vocabolari controllati, che ora possono essere riutilizzati nel contesto di applicazioni innovative.

L'adattamento di un sistema di rappresentazione della conoscenza esistente (ad esempio i tesauri, le liste di intestazione per soggetto o uno schema di classificazione) al formato SKOS richiede la conversione e la marcatura dei dati in un vocabolario XML specifico. Questa attività può essere automatizzata, e le istituzioni possono facilmente applicare tecniche di elaborazione del testo (ad esempio tramite linguaggio di *scripting* in Python o Perl) per aggiungere i tag necessari ai vocabolari esistenti. Ma lo sviluppo di un progetto LOD non richiede soltanto la conversione dei dati in RDF o in schemi a questo collegati. È necessario anche dare alle persone la possibilità di utilizzare i nostri dati, e nel dettaglio, oppure di utilizzare tali dati tramite applicazioni software di terze parti. Si tratta chiaramente di una delle sfide più importanti e sottovalutate dei progetti relativi al Web semantico e ai LOD. La finalità di tali progetti non dovrebbe limitarsi alla pubblicazione dei dati in un modo diverso, utilizzando formati diversi, per un pubblico umano. L'obiettivo delle iniziative legate al Web semantico è quello di supportare un riuso efficiente dei dati da parte di altre applicazioni software.

Per raggiungere tale obiettivo, le istituzioni devono considerare di dedicare un impegno significativo alla possibilità di rendere disponibili i dati, in modo da consentire uno sfruttamento dinamico. Allo stato attuale, la maggior parte delle iniziative LOD sviluppate da biblioteche e archivi sono focalizzate sulla conversione di dati esistenti e sulla loro pubblicazione in formato RDF. Ma queste attività non daran-

RICARDO EITO-BRUN, Universidad Carlos III de Madrid, Madrid, España, e-mail reito@bib.uc3m.es.

¹ World Wide Web Consortium (W₃C), *SKOS Simple knowledge organization system: W₃C recommendation 18 August 2009*, <<https://www.w3.org/TR/skos-reference/>>.

no i risultati attesi, a meno che un riuso efficiente di tali dati venga supportato da ulteriori tecnologie e strumenti.

Tecnologie per la diffusione dei dati

Una volta che i dati sono convertiti in RDF, o in formati collegati a RDF (ad esempio SKOS), le organizzazioni devono fornire strumenti per interrogare e utilizzare tali dati. Nella varietà di tecnologie e di standard del Web semantico è presente anche il linguaggio SPARQL, che è simile al linguaggio SQL utilizzato per interrogare basi di dati relazionali. Per interrogare un *dataset* in RDF utilizzando questo linguaggio di ricerca, una pratica comune è quella di stabilire punti di arrivo o estremi in SPARQL che forniscano URL verso i quali possano essere dirette le ricerche in SPARQL all'interno di un *dataset*, e ricavare la lista dei risultati ottenuti in formati diversi (di solito in XML). Gli estremi in SPARQL vengono costruiti sulla base di 'archivi semantici' che immagazzinano i dati in RDF o SKOS. Tali archivi semantici offrono una gamma di funzionalità tipiche, fra le quali l'importazione in blocco di dati in RDF e la disponibilità di estremi in SPARQL che possono accettare, analizzare ed eseguire interrogazioni in SPARQL e restituirne i risultati.

I problemi principali nell'utilizzo di estremi in SPARQL e di archivi semantici sono: a) l'assenza di integrazione di tali strumenti con strumenti che possano editare dati in RDF o SKOS (i file di input devono essere editati con specifici strumenti prima di essere caricati nell'archivio); b) l'accesso ai dati è riservato agli utilizzatori del linguaggio di interrogazione in SPARQL; c) non tutti i dati gestiti dalle biblioteche e dagli archivi sono codificati in RDF o in formati che si basano su RDF, ad esempio, possono essere utilizzati altri formati che si basano su XML, come i record di autorità in EAC (Encoded archival context). Il problema principale di un approccio basato su estremi SPARQL è che questi sono difficili da utilizzare, perché gli utenti devono conoscere la sintassi di SPARQL, piuttosto complessa, per integrare le applicazioni destinate ai 'consumatori'.

Nel caso di studio analizzato in questo contributo, tali ragioni hanno condotto alla selezione di uno strumento XML nativo per costruire e fornire accesso a una base dati ricavata da dati codificati in SKOS.

Conversione e caricamento di una base dati in SKOS

Il progetto qui esaminato aveva lo scopo di pubblicare due tesauri già creati in passato dal Ministero dello sviluppo spagnolo². I tesauri coprivano aree diverse nell'area dell'ingegneria civile. Inizialmente dovevano servire a mappare i concetti utilizzati per indicizzare le descrizioni delle opere di ingegneria civile e documentare le pratiche. Tali descrizioni vennero create con lo standard EAD (Encoded archival description). Ai catalogatori addetti alla descrizione dei materiali venne fornito uno strumento online, grazie al quale i termini appropriati potevano essere ricercati e selezionati all'interno dei due tesauri citati. L'accesso ai tesauri era incorporato in una funzionalità di ricerca aggiunta all'editor XML utilizzato per la scrittura dei record in EAD.

Per assicurare tale funzionalità, i tesauri esistenti, in formato cartaceo, vennero digitalizzati e convertiti in formato elettronico. Utilizzando una serie di comandi automatizzati, venne poi completato il processo di inserimento dei tag SKOS XML all'interno del contenuto dei tesauri, grazie all'adozione delle etichette originali dei tesauri per i termini ampi (BT, 'broad terms'), ristretti (NT, 'narrow terms') o corre-

² Vedere: <http://www.cehopu.cedex.es/etm/etm_index.htm>.

lati (RT, 'related terms'), oltre che sostituendoli con i corrispondenti tag in SKOS. Dopo aver codificato i tag secondo le specifiche SKOS, i tesauri sono stati caricati in una base dati in XML per fornire funzionalità di interrogazione e ricerca ad agenti esterni (sia utilizzatori umani che applicazioni informatiche).

In una seconda fase, i tesauri, ormai taggati, sono stati suddivisi in file separati, mediante un foglio di stile XSLT con l'ausilio dello strumento Altova® MapForce™. Per ogni concetto presente nei tesauri sono stati creati file XML distinti. Ogni file conteneva i dati chiave su ogni concetto: forma autorizzata, alternativa, forme non autorizzate, termini ampi, ristretti e correlati. Per rendere eseguibili i collegamenti fra i diversi concetti nei tesauri, a ogni concetto è stato attribuito un URI (Uniform resource identifier). I collegamenti e i riferimenti incrociati fra i termini ampi, ristretti e correlati sono stati costruiti utilizzando gli URI dei termini ricercati. In questo modo, è stato possibile mantenere le relazioni fra i concetti, nonostante fossero archiviati in file separati.

Infine, i vari file sono stati caricati in una base di dati XML nativa. Si è deciso di utilizzare, come strumento di archiviazione dei dati, Oracle® Berkeley DBXML. Questa base di dati corrisponde agli strumenti XML integrati in basi di dati che assolvono bene al proprio compito e forniscono una buona affidabilità, minimizzando i compiti di tipo amministrativo. Anche se il Berkeley DBXML non supporta il linguaggio di interrogazione in SPARQL, offre un linguaggio di interrogazione potente che utilizza i tag XML e può interagire con la base di dati tramite le API (Application programming interfaces), disponibili per vari linguaggi di programmazione: JAVA, PHP, C, ecc.

In questo progetto, l'interazione e la comunicazione fra lo strumento esterno (l'editor EAD/ISAD(G) XML) e la base di dati Berkeley DBXML sono state effettuate tramite servizi web REST, implementati con PHP. Per standardizzare i meccanismi di interrogazione, l'interazione e lo scambio di messaggi fra la base di dati, lo strumento per l'utenza finale è stato costruito utilizzando il protocollo tecnico SRU pubblicato dalla Library of Congress.

Grazie all'utilizzo di questi servizi e dei messaggi così come definiti nel protocollo tecnico SRU, gli utenti sono stati in grado di interrogare e interagire con i tesauri digitalizzati. Sono stati messi nelle condizioni di cercare, esplorare e navigare nell'intera rete dei concetti. Le interrogazioni supportavano anche varie possibilità di scelta: ad esempio, la ricerca dei soli termini preferiti, la ricerca di un qualsiasi termine nel vocabolario, la ricerca di termini con un significato più specifico o più generico rispetto a un altro, eccetera.

Conclusioni

Negli ultimi anni, è aumentato l'utilizzo di tecnologie semantiche per diffondere e aprire i dati a terze parti e sostenere lo sviluppo di applicazioni innovative e distribuite. Sempre più organizzazioni vengono coinvolte in questo tipo di progetti, e il Web semantico, insieme ai LOD, stanno offrendo nuove opportunità per sfruttare il valore delle basi di dati esistenti.

La pubblicazione di tesauri – o di altre basi di dati – richiede un processo a due fasi: la conversione di dati esistenti in un formato adeguato, e il ricorso a strumenti di interrogazione dei dati. Questo contributo ha descritto l'utilizzo combinato di SKOS e di una base di dati integrata XML nativa, per creare un punto di accesso a un tesoro digitalizzato. L'approccio ha dimostrato di essere adeguato per lo sviluppo di una infrastruttura tecnica per la gestione di contenuti e metadati delle biblioteche del Ministero dello sviluppo spagnolo.

ABSTRACT

RICARDO EITO-BRUN, Universidad Carlos III de Madrid, Madrid, España, e-mail reito@bib.uc3m.es.

Varie istituzioni e organizzazioni hanno cominciato a porre attenzione alle tecnologie del Web semantico quale componente essenziale della propria infrastruttura per la gestione delle informazioni. Un aspetto chiave nei sistemi informativi tradizionali che avrà una rilevanza sostanziale nel Web semantico e nei progetti di linked open data sono i sistemi di organizzazione della conoscenza: le classificazioni, i tesauri, le liste di intestazione per soggetto eccetera. Gli standard tecnici W3C per il Web semantico includono la specifica SKOS (Simple knowledge organization system) per codificare e taggare tali strumenti di rappresentazione delle informazioni. L'adattamento di un sistema di rappresentazione della conoscenza alla specifica SKOS richiede la conversione e la marcatura dei dati.

Modelli per la comunicazione dei dati di ricerca in archeologia: il caso dei Getty vocabularies come linked open data

di Angela Bellia e Fiammetta Sabba

I linked data per i beni culturali

Il modello dei linked open data (LOD) ampiamente consolidato per i dati governativi e scientifici sta assumendo un ruolo centrale nell'organizzazione e condivisione dei dati sul Web anche nel mondo dei beni culturali. Tale affermazione emerge sia nel dibattito italiano sia in quello internazionale per l'importanza assunta dal tema dei LOD nella promozione del patrimonio culturale attraverso la rete e per la capacità del modello di far convergere pratiche descrittive tradizionalmente distanti. I linked data creano le premesse per un accesso unificato e globale, e soprattutto qualificato e controllato, alle risorse culturali, mettendole a disposizione di tutti. L'ambiente di informazione aperto costruito dal modello LOD ha un enorme potenziale per affrontare alcune questioni connesse al campo dei beni culturali, tra cui l'interoperabilità dei dati e la loro integrazione. Attraverso un approccio di relazioni tra i dati espliciti e trattabili da un punto di vista computazionale sono così possibili nuove forme di accesso all'informazione e nuove opportunità per l'interpretazione di dati e documenti¹. All'interno di una cornice tecnologica condivisa e di una nuova rappresentazione i dati e i metadati di biblioteche, archivi e musei sono collegati attraverso una metodologia che ha lo scopo di crearli, condividerli e riutilizzarli. I linked data sono così una delle strategie tecnologiche più avanzate in favore della gestione, catalogazione, valorizzazione e comunicazione di tutte le tipologie di beni culturali che richiedono non soltanto una radicale revisione dell'approccio alla rappresentazione tradizionale dei dati, alla loro conservazione e ricerca, ma anche la necessità di formare figure che, al di là dello specifico ambito di azione (bibliotecario, archivistico, museale), abbiano un profilo ben definito sul piano dell'aggiornamento e della contemporaneità.

In Italia si stanno diffondendo alcune iniziative che mirano alla formazione di figure che sappiano combinare la preparazione culturale con quella tecnica nel trattamento informativo e comunicativo dei beni culturali. Tra queste iniziative si trova la *summer school* Linked data per i beni culturali, diretta da Fiammetta Sabba, che si

ANGELA BELLIA, New York University, Institute of fine arts, New York, United States of America; Alma mater studiorum – Università di Bologna, Dipartimento di beni culturali, Ravenna, e-mail angela.bellia@unibo.it.

FIAMMETTA SABBA, Alma mater studiorum – Università di Bologna, Dipartimento di beni culturali, Ravenna, e-mail fiammetta.sabba@unibo.it.

Ultima consultazione siti web: 1 dicembre 2016.

1 Hilary K. Thorsen; M. Cristina Pattuelli, *Linked open data and the cultural heritage landscape*. In: *Linked data for cultural heritage*, edited by Ed Jones, Michele Seikel. London: Facet, 2016, p. 1-22.

è svolta presso il Dipartimento di beni culturali dell'Università di Bologna, Sede di Ravenna². La *summer school* è stata un momento formativo prezioso perché ha consentito ai partecipanti di conoscere i principi e le tecniche dei linked data applicati ai beni culturali³.

Il progetto Linked jazz

Uno dei progetti analizzati nel corso della *summer school* è stato il Linked jazz and the networked catalog at New York Pratt Institute Library: data modeling con creazione di triple da testo e da metadati. Il progetto è stato creato allo scopo di usare la tecnologia dei linked data per scoprire connessioni significative tra documenti e dati relativi ai protagonisti e ai professionisti della musica jazz e di sviluppare i relativi strumenti e metodi dei linked data. Il Linked jazz ha mostrato come l'applicazione della tecnologia LOD alle risorse dei beni culturali, in questo caso musicali, sia stata ottenuta nella fase di avvio filtrando un elenco di nomi di musicisti jazz (circa 9.300) estratti da DBPedia, la versione LOD di Wikipedia, e anche attraverso Europea e la Digital Public Library of America (DPLA). Uno dei risultati più significativi emersi dall'incrocio e dalla condivisione dei dati composti da triple rappresentanti discografia, fotografie di musicisti, *performances*, patrimonio librario, ha evidenziato la presenza e l'attività di numerose ottime musiciste jazz che hanno avuto una forte influenza nella formazione di molti colleghi maschi più famosi. Grazie alla tecnologia LOD, che consente di espandere l'accesso tradizionale ai contenuti d'archivio, la memoria di queste musiciste non soltanto non è andata perduta ma è stata resa disponibile in un ambiente di ricerca globale e integrato.

Un caso di studio: i Getty vocabularies

Tra i dati del patrimonio culturale meno visibili vi sono stati quelli appartenenti ai musei che, per essere pubblicati on-line, necessitano di licenze e informazioni sulla provenienza. A questo si aggiunge che per i musei è estremamente importante che i dati pubblicati siano esatti per mantenere la credibilità e l'affidabilità dell'istituzioni museali all'interno del mondo scientifico. Tuttavia, recentemente i musei hanno iniziato ad utilizzare la tecnologia LOD per rendere visibili i loro dati e le loro collezioni sul Web. Un interessante esempio sono i Getty vocabularies as linked open data del Getty Research Institute di Los Angeles⁴. Si tratta di un caso di studio che mira a esplorare modelli innovativi per migliorare la scoperta e l'interpretazione del patrimonio culturale attraverso l'applicazione della tecnologia LOD per archivi digitali. Questo ampio progetto è parte del continuo impegno del Getty a mettere le risorse del Getty Research Institute liberamente disponibili a tutti. Con l'obiettivo gene-

2 Questa attività rientrava nell'ambito del progetto di ricerca TELESTES, finanziato dal programma europeo Marie Curie Actions. International outgoing fellowships for career development presso l'Institute of Fine Arts della New York University (outgoing) e il Dipartimento di beni culturali dell'Università di Bologna (ingoing): <http://cordis.europa.eu/project/rcn/185858_en.html>. La *summer school* si è svolta a Ravenna dal 16 al 20 maggio 2016: <<https://eventi.unibo.it/linked-data-per-i-beni-culturali>>. Il programma della prossima edizione è disponibile all'indirizzo: <<https://eventi.unibo.it/linked-data-per-i-beni-culturali/news>>.

3 Oltre a un'introduzione teorica e pratica ai più recenti metodi di rappresentazione delle risorse conservate in biblioteche, archivi e altri tipi di istituzioni culturali, la *summer school* ha trattato i metodi e le tecniche per accrescerne la visibilità e facilitarne l'accesso.

4 <<http://www.getty.edu/research/tools/vocabularies/lod/index.html>>.

rale di contribuire a rendere visibile la ricca e diversificata rete di connessioni, il progetto Getty vocabularies as linked open data è un significativo esempio di pratiche di creazione LOD che contribuisce alla crescente ricerca di LOD in biblioteche, archivi e musei. I Getty vocabularies sono stati costruiti per permettere di classificare, descrivere e indicizzare gli oggetti e le informazioni relative al patrimonio culturale: con una terminologia multilingue sono compresi arti visive, architettura, dati sulla conservazione, beni archeologici, materiali d’archivio e materiale bibliografico. Inoltre, essi sono conformi alle norme internazionali e si sviluppano attraverso contributi che la tecnologia trasforma in un set di dati leggibili, incorporandoli progressivamente nel Web semantico: pubblicando i Getty vocabularies come LOD, il Getty Research Institute condivide i risultati di oltre trenta anni di studi e ricerche.

Per supportare lo scambio, l’uso e il riuso dei metadata, i Getty vocabularies adottano un linguaggio semplice per formalizzare i concetti espressi in formato semantico, utilizzando RDF e ontologie adeguate che condividono la medesima rappresentazione semantica di base: essi includono l’*Art and architecture thesaurus* (AAT) (figura 1), il *Thesaurus of geographic names* (TGN) e l’*Union list of artist names* (ULAN), e contengono una terminologia strutturata per l’archeologia, l’arte, l’architettura, le arti decorative, i materiali d’archivio e il materiale bibliografico, che cresce e si evolve attraverso contributi in ambienti connessi e aperti, offrendo un potente canale per la ricerca. I dati sono conformi agli standard internazionali e forniscono informazioni autorevoli non soltanto per i catalogatori, ma anche e soprattutto per i ricercatori e i fornitori di dati. Nel *Getty vocabulary semantic representation* è disponibile anche una descrizione del set completo di URI.

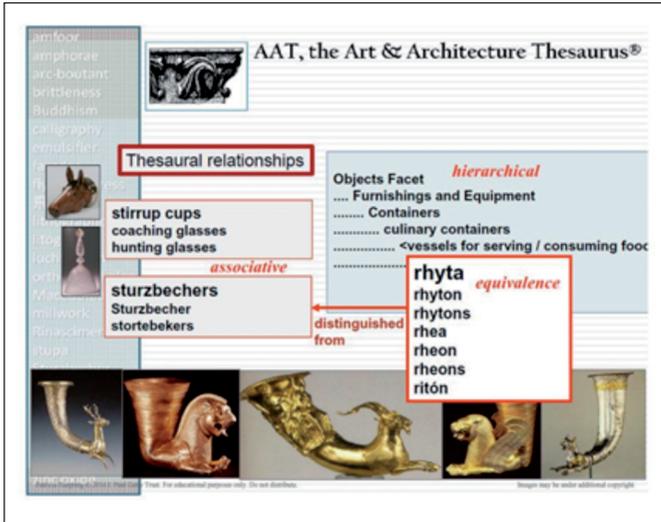


Figura 1 – Relazioni nell’*Art and architecture thesaurus*

Modelli per la comunicazione nella ricerca archeologica

Dall’analisi di questo caso di studio sono scaturite alcune riflessioni riguardanti la ricerca archeologica. Mettere i Getty vocabularies a disposizione della comunità di ricerca come LOD può aver un effetto di trasformazione e di evoluzione nel campo dell’archeologia, e delle discipline a essa strettamente correlate, e nei suoi metodi di

indagine perché contribuisce a mantenere la piena collaborazione della comunità LOD, anche attraverso un forum di discussione pubblica.

L'emergere di un nuovo rapporto tra dati digitali e metodologia archeologica, che non è nato recentemente, si inserisce all'interno di un più ampio fenomeno che mira alla connessione di una catena di relazioni e interazioni virtuali, e contribuisce alla creazione di differenti forme comunicative e totalmente digitali. Nel settore dell'antichistica, e in particolare in quello archeologico, queste soluzioni hanno l'obiettivo di condividere le informazioni a un livello mai conosciuto prima. Non si tratta di una semplice espansione quantitativa dei dati disponibili sugli archivi consultabili, ma di una metodologia nuova che tiene conto dell'esigenza di rendere le informazioni comprensibili sia alle macchine che agli utenti.

Dall'analisi dell'esperienza del Getty Research Institute emerge come un tale approccio possa contribuire a creare nuovi modelli e pratiche per la pubblicazione di dati in campo archeologico con l'obiettivo di alimentare lo studio collaborativo, l'analisi e, soprattutto, la condivisione delle informazioni che mirano ad una nuova trasmissione delle conoscenze. Il progetto del Getty, inoltre, sembra mostrare quel salto di qualità (e non soltanto di quantità) nello studio dei dati, nell'aggregazione delle informazioni e infine nel loro corretto riuso. Se da un lato queste caratteristiche del progetto risultano utili e indispensabili al rinnovamento della metodologia archeologica, e dei connessi campi di indagine, dall'altro mettono a disposizione strumenti e formati adatti a favorire e accrescere la condivisione in rete con nuovi metodi per la concettualizzazione e l'organizzazione della conoscenza nel più ampio spazio dell'approccio collaborativo online.

ABSTRACT

ANGELA BELLIA, New York University, Institute of fine arts, New York, United States of America; Alma mater studiorum – Università di Bologna, Dipartimento di beni culturali, Ravenna, e-mail angela.bellia@unibo.it.

FIAMMETTA SABBA, Alma mater studiorum – Università di Bologna, Dipartimento di beni culturali, Ravenna, e-mail fiammetta.sabba@unibo.it.

Questo articolo presenta i metodi e gli strumenti creati nell'ambito del progetto Getty vocabularies as linked open data. Il progetto esplora modelli innovativi per migliorare la scoperta e l'interpretazione del patrimonio culturale attraverso l'applicazione della tecnologia linked open data (LOD) per archivi digitali. Esso è parte del continuo impegno di rendere le risorse del Getty Research Institute liberamente disponibile a tutti. Con l'obiettivo generale di contribuire a rendere visibile la ricca e diversificata rete di connessioni, il progetto Getty vocabularies è un caso di studio di pratiche di creazione LOD e contribuisce alla crescente ricerca sui LOD in biblioteche, archivi e musei. Viene discusso come i Getty vocabularies possano contribuire a creare nuovi modelli e pratiche per la pubblicazione di dati in campo archeologico al fine di alimentare lo studio collaborativo, l'analisi e la condivisione delle informazioni.

Coordinamento delle biblioteche speciali e specialistiche di Torino (CoBiS): Progetto linked open data

di Federico Morando, Emanuela Secinaro e il Gruppo CoBiS LOD

Il CoBiS e il Progetto LOD

Il CoBiS è il Coordinamento delle biblioteche speciali e specialistiche dell'area metropolitana torinese¹. È stato istituito a giugno 2008 da una decina di biblioteche, con lo scopo di creare una rete condivisa per l'aggiornamento professionale e il miglioramento del servizio all'utenza. Si è progressivamente ampliato e ora è costituito da 65 biblioteche con i relativi patrimoni librari e archivistici.

Gli utenti delle biblioteche CoBiS sono principalmente mossi dall'interesse verso la specificità dei documenti posseduti e spesso necessitano di aiuto qualificato nella consultazione e nella ricerca bibliografica. L'esigenza di formazione dei bibliotecari e di valorizzazione dei singoli patrimoni, rendendoli conosciuti e accessibili, ha dato origine al Progetto linked open data (LOD) del CoBiS. Al momento è l'unico sui LOD in Piemonte e può essere considerato il progetto pilota della Regione. L'obiettivo futuro è coinvolgere anche le altre biblioteche del territorio afferenti alle diverse istituzioni.

Genesi del Progetto linked open data

Il Progetto LOD è iniziato con un percorso formativo in collaborazione con Maurizio Vivarelli dell'Università di Torino e si è articolato in vari appuntamenti collegati: diritto d'autore, collaborazione tra biblioteche e Wikipedia, open data.

Si è poi sviluppato con una riflessione sui LOD applicati ai dati catalografici. Sette biblioteche del CoBiS si sono proposte per la sperimentazione, con la collaborazione scientifica del Centro Nexa su Internet & società del Politecnico di Torino e il supporto tecnologico di Synapta Srl: Accademia delle scienze di Torino, Associazione archivio storico Olivetti, Biblioteca nazionale del club alpino italiano, INRIM – Istituto nazionale di ricerca metrologica, Deputazione subalpina di storia patria, Educatorio della provvidenza, INAF – Osservatorio astrofisico di Torino.

FEDERICO MORANDO, Synapta srl; Politecnico di Torino, Nexta center for internet & society, Torino, e-mail federico.morando@gmail.com.

EMANUELA SECINARO, Istituto nazionale di ricerca metrologica, Biblioteca, Torino, e-mail e.secinaro@inrim.it.

Il Gruppo CoBiS LOD è composto da: Barbara Bonino, Elena Borgi, Maria Pia Girelli, Gabriella Morabito, Federico Morando, Emanuela Secinaro, Luisa Schiavone, Anna Maria Viotto.

Ultima consultazione siti web: 30 novembre 2016.

¹ <www.cobis.to.it>.

L'eterogeneità delle fonti: i dati in gioco

Di seguito una breve presentazione delle biblioteche e dei dati messi a disposizione per la sperimentazione.

La Biblioteca dell'Istituto nazionale di ricerca metrologica (INRIM)² ha estratto dal proprio catalogo le 1.600 monografie registrate. Questa biblioteca utilizza per la catalogazione il software ErasmoNet³ e i dati sono stati estratti generando una tabella Excel.

La Biblioteca dell'Osservatorio astrofisico di Torino⁴ fa parte dell'Istituto nazionale di astrofisica – INAF⁵. I cataloghi delle biblioteche INAF sono confluiti in un catalogo unico⁶ gestito con il software BIBLIOWin 5.oweb⁷. Per il progetto, i dati bibliografici vengono esportati in XML in tempo reale, selezionando le notizie relative all'Osservatorio astrofisico di Torino⁸.

La biblioteca dell'Accademia delle scienze⁹ ha un patrimonio stimato di quasi 300.000 volumi, di cui oltre il 15% antichi, gestito con Sebina open library e inserito nel Polo bibliografico della ricerca che fa capo all'Università di Torino. Il patrimonio archivistico è catalogato con xDams. Sia Sebina che xDams prevedono la possibilità di *export* in XML.

La Biblioteca dell'Associazione archivio storico Olivetti (AASO) raccoglie le monografie e i 187 periodici italiani ed esteri editati dalla Società Olivetti S.p.A. La Biblioteca fa parte del Sistema bibliotecario di Ivrea e Canavese e ha adottato per la catalogazione il software ErasmoNet. Per il Progetto LOD mette a disposizione i 1.800 titoli, in gran parte di letteratura grigia, del fondo Centro formazione meccanici (CFM).

La Biblioteca nazionale del club alpino italiano¹⁰, a Torino dal 1863, aderisce a SBN dal 2004 e dal 2014 coordina il Catalogo unico dei beni culturali del Museo nazionale della montagna (7.500 notizie) e delle Sezioni CAI (33.000 notizie) gestito col software Clavis.

La Biblioteca della Deputazione subalpina di storia patria¹¹, mette a disposizione per il progetto i dati catalografici delle edizioni della Deputazione subalpina, che iniziano nel 1836 con la pubblicazione del primo volume degli *Historiae patriae monumenta*, e proseguono, sino a oggi, con varie collane, come quella della Biblioteca di storia italiana recente, degli Studi e documenti di storia economica, degli atti dei congressi.

La biblioteca storica dell'Educatario della provvidenza¹² si compone di due settori: libri e fotografie. La sezione fotografica comprende 716 diapositive degli anni

2 <<http://rime.inrim.it/biblioteca/>>

3 <<http://www.cs.erasmo.it/soluzioni/servizi-beni-culturali/erasmonet/>>.

4 <www.oato.inaf.it>.

5 <www.inaf.it>.

6 <www.bibliowin.it/inaf/>.

7 <<http://www.cgsi.it>>.

8 In questa prima fase del progetto, il cui finanziamento è affidato per metà alla Regione Piemonte, si intende dare maggior rilievo al patrimonio locale, senza tuttavia escludere un coinvolgimento più ampio a livello territoriale e istituzionale.

9 <http://www.accademiadellesienze.it/biblioteca_e_archivio>.

10 <<http://mnmt.comperio.it/biblioteche-cai/Biblioteca-Nazionale/>>.

11 <<http://www.deputazionealpina.it/biblioteca.html>>.

12 <http://www.cobis.to.it/Web/Biblioteche_aderenti/Biblioteca_storica_dell_Educatario_della_Provvidenza_Paolo_Girelli_e_Vera_Tua>.

Trenta di storia dell'arte e fotografie del periodo compreso tra la fine dell'Ottocento e gli anni Settanta del Novecento. Da questa raccolta sono state tratte alcune foto e relativi metadati.

Aspetti tecnici del Progetto LOD del CoBiS

Il Progetto LOD mira a fornire il CoBiS di un'infrastruttura per pubblicare LOD, creando una *pipeline* di triplificazione (trasformazione di dati in linked data) facilmente automatizzabile e replicabile. Tutto ciò utilizzando tecnologie *open source*, tra le quali il linguaggio di mappatura RML¹³.

Il linguaggio di mappatura RML

L'RDF mapping language (RML) è un linguaggio generico, che può essere utilizzato per esprimere (utilizzando triplette RDF) le regole per trasformare dati eterogenei (XML, JSON o dati tabulari) in linked data.

RML è disponibile liberamente con licenza MIT su GitHub¹⁴ e Synapta ha apportato alcune migliorie all'RML-processor, per velocizzare (di circa due ordini di grandezza) il processo di triplificazione.

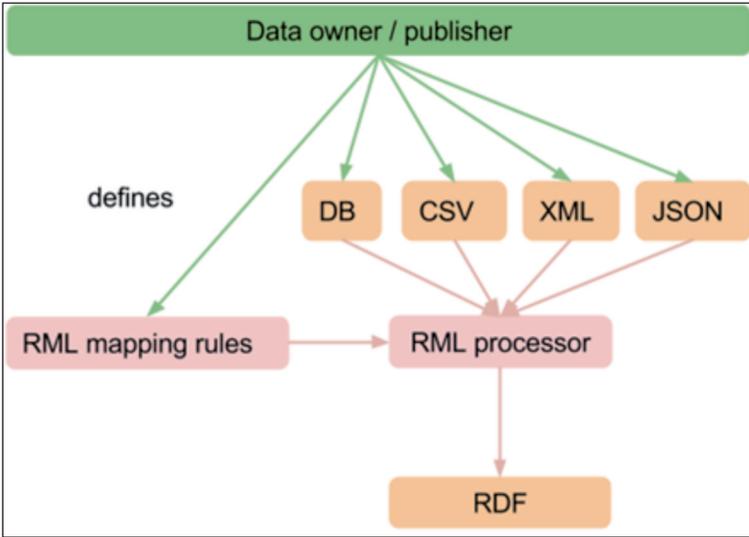


Figura 1¹⁵

Architettura

RML è solo una delle componenti dell'architettura software complessiva del Progetto LOD del CoBiS, rappresentata in figura 2.

13 <<http://rml.io/>>, realizzato dal gruppo iMinds dell'Università di Gand.

14 <<https://github.com/RMLio>>.

15 Fonte: <<http://rml.io/>>, © 2013–2016 Ghent University – iMinds.

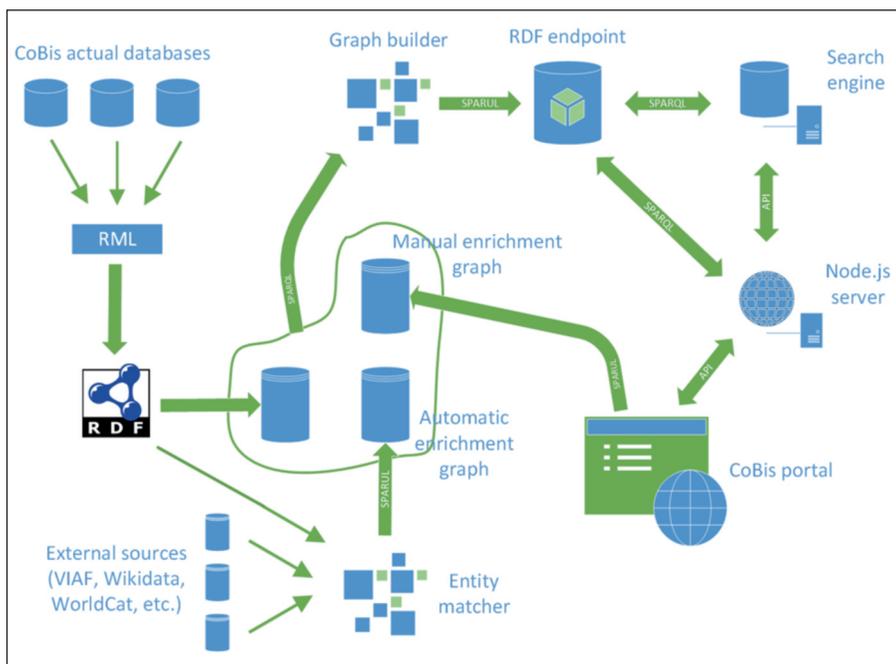


Figura 2

I dati eterogenei delle banche dati del CoBiS vengono processati sulla base delle map-pature RML, per produrre un primo grafo di LOD ‘grezzi’ (ovvero, espressi in RDF¹⁶, ma non collegati ad altre fonti), che viene salvato in un apposito database (*triplestore*). In parallelo, vengono raccolti una serie di dati da fonti esterne (VIAF, Wikidata, WorldCat ecc.). Il componente denominato *Entity matcher* effettua un primo inter-linking automatico dei LOD ‘grezzi’. Eventuali regole manuali di gestione delle relazioni tra le diverse fonti vengono salvate in un altro grafo di servizio, che potrà poi essere aggiornato dal *back-end* del portale CoBiS (ad esempio, i bibliotecari potranno aggiungere o correggere manualmente delle relazioni di identità tra autori e/o opere censiti da fonti differenti). Un altro elemento, il *Graph builder*, si farà dunque carico di creare gli URI di riferimento per il CoBiS e di scrivere su un *triplestore* di riferimento la versione ‘di sintesi’ dei dati raccolti. Un motore di ricerca indicizzerà questi dati, per una fruizione più efficiente dal portale CoBiS (alimentato da un server Node.js, sulla base del motore di ricerca e di query SPARQL live).

Tutta l’architettura è basata su tecnologie *open source*¹⁷ ed il software sviluppato da Synapta per il Progetto LOD del CoBiS sarà rilasciato con licenze *open source*.

¹⁶ Il formalismo RDF (Resource description framework) è lo standard per la rappresentazione dei linked open data: <<https://www.w3.org/RDF/>>.

¹⁷ Saranno *open source* tutte le componenti principali, per facilitare il riutilizzo; si sta valutando, per ottimizzare costi e *performance*, l’utilizzo di alcune tecnologie proprietarie, ad esempio SOPHIA semantic search di CELI per il motore di ricerca.

Interlinking e ontologie di riferimento

Buona parte dell'attività di *interlinking* automatico sarà volta alla connessione delle informazioni relative alle collezioni del CoBiS al resto del mondo, per arricchire le informazioni catalografiche di dati aggiuntivi. Ma anche l'*interlinking* inverso – ovvero, l'aggiunta di collegamenti verso le opere del CoBiS su piattaforme terze – fa parte del progetto, con lo scopo di rendere maggiormente visibili online le biblioteche del Coordinamento.

Per le ontologie di riferimento, il progetto è improntato al riutilizzo degli standard più diffusi, in particolare la Bibliographic framework initiative (BIBFRAME) avviata dalla Library of Congress, ma anche Schema.org, per massimizzare l'impatto sui motori di ricerca.

Interfaccia utente

Il principale modello del Progetto LOD è rappresentato dalla 'semplicità' (almeno a livello di apparenza estetica) dell'interfaccia di WorldCat.org, ma progetti più focalizzati sui linked data, a partire dall'esperienza SHARE Catalogue¹⁸, offrono ulteriori spunti.

Note conclusive e sviluppi futuri

La *roadmap* progettuale prevede che il pilota – che rappresenta la prima componente di una più ambiziosa CoBiS digital library – sia online in versione beta entro la primavera 2017¹⁹.

ABSTRACT

FEDERICO MORANDO, Synapta s.r.l.; Politecnico di Torino, Nexta center for internet & society, Torino, e-mail federico.morando@gmail.com.

EMANUELA SECINARO, Istituto nazionale di ricerca metrologica, Biblioteca, Torino, e-mail e.secinaro@inrim.it.

Il Gruppo CoBiS LOD è composto da: Barbara Bonino, Elena Borgi, Maria Pia Girelli, Gabriella Morabito, Federico Morando, Emanuela Secinaro, Luisa Schiavone, Anna Maria Viotto.

Il CoBiS è una rete di 65 biblioteche specialistiche. Il Progetto linked open data (LOD) del CoBiS è il pilota del Piemonte per pubblicare LOD, utilizzando tecnologie *open source*, quali il linguaggio di mappatura RML.

La prima sfida del progetto consiste nel riconciliare dati provenienti da cinque diversi applicativi (Clavis, ErasmoNet, BIBLIOWin, SbnWeb e Sebina OpenLibrary), linkando altresì contenuti multimediali e alcuni dati archivistici.

L'*interlinking* si focalizza su VIAF e Wikidata e le ontologie di riferimento sono BIBFRAME e Schema.org.

¹⁸ <<http://catalogo.share-cat.unina.it/>>.

¹⁹ All'indirizzo <<http://dati.cobis.to.it>>.

Gli *authority data* per l'integrazione *cross-domain* dei beni culturali: riflessioni su un approccio alla lettura trasversale dei beni culturali della Chiesa cattolica italiana

a cura del Gruppo di lavoro sugli *authority file* dell'Ufficio per i beni culturali ecclesiastici e l'edilizia di culto della Conferenza episcopale italiana

L'obiettivo del contributo è presentare la metodologia di lavoro per la gestione e l'utilizzo degli *authority file* all'interno del portale di aggregazione Beweb – beni ecclesiastici in Web, dove i dati di autorità fungono da punto di snodo.

Il contesto di riferimento del portale Beweb aiuta a comprendere meglio le soluzioni individuate¹.

Il portale Beweb – Beni ecclesiastici in Web

Beweb è il portale integrato dei beni culturali ecclesiastici, che raccoglie e presenta il lavoro di catalogazione e censimento sistematico di tutti i beni culturali ecclesiastici – iniziato nel 1996 e tuttora in corso – portato avanti dalle diocesi italiane e dagli istituti culturali ecclesiastici sui beni di loro proprietà.

In Beweb convivono insieme beni storico artistici, architettonici, archivistici e librari, con banche dati in continuo aggiornamento e incremento. Ma sono presenti anche altri 'oggetti', che hanno l'obiettivo di fornire degli strumenti di avvicinamento ai contenuti del portale. Tra questi le voci di glossario, che sono disponibili sia nella home page generale che in quelle di settore; le pagine descrittive delle diocesi e delle regioni ecclesiastiche, dove si trovano informazioni storiche e di utilità su queste istituzioni territoriali ecclesiastiche; le pagine degli istituti culturali ecclesiastici, che valorizzano i musei, le biblioteche e gli archivi censiti nell'Anagrafe degli istituti culturali ecclesiastici della CEI. E sono presenti anche schede di autorità persona, famiglia, ente.

I progetti di censimento e catalogazione sono promossi dall'Ufficio per i beni culturali ecclesiastici e l'edilizia di culto della CEI adottando un modello di lavoro distribuito: il lavoro di catalogazione vero e proprio viene svolto dai singoli operatori sul territorio nazionale (circa 3.000 tra bibliotecari, archivisti, storici dell'arte e architetti, unici garanti della scientificità del lavoro), mentre il gruppo di lavoro sui beni culturali si occupa a livello centrale del trattamento di questi dati e della loro presentazione, valorizzazione e fruizione sul Web.

La complessità e la varietà dei beni culturali oggetto delle campagne di censi-

Gruppo di lavoro sugli *authority file* dell'Ufficio per i beni culturali ecclesiastici e l'edilizia di culto della Conferenza episcopale italiana, composto da Adriano Belfiore, Francesca Maria D'Agnelli, Laura Gavazzi, Claudia Guerrieri, Maria Teresa Rizzo, Silvia Tichetti, Paul Gabriele Weston, sito web <<http://www.beweb.chiesacattolica.it/it>>, e-mail beweb@chiesacattolica.it.

1 <www.chiesacattolica.it/beweb>.

mento e catalogazione è riconoscibile già dalla home page, dove troviamo un menu articolato a seconda delle diverse specificità degli ambiti culturali presentati: beni storico-artistici, beni architettonici, beni librari, beni archivistici e istituti culturali.

Le pagine web dedicate al singolo settore propongono delle modalità di ricerca e lettura proprie del singolo ambito catalografico: ad esempio nella pagina beni storico-artistici o in quella dedicata agli edifici di culto si nota la possibilità di cercare per cronologia, sulla pagina beni librari si trova in evidenza la ricerca per soggetto e in quella dei beni archivistici si può cercare per soggetto produttore o conservatore.

Ma la novità di Beweb è quella dell'approccio trasversale alla lettura dei beni culturali, visibile in home page, dove ad esempio a partire da una ricerca semplice è possibile interrogare basi dati di ambito diverso e ottenere risultati variegati ma con risorse sempre relazionate tra loro.

La composizione di Beweb e le sue potenzialità narrative si evidenziano apprezzando la circolarità della navigazione attraverso schede di autorità. A partire dalla home page di Beweb possiamo cercare informazioni generiche su una persona, ad esempio l'architetto Elia Fornoni.

I risultati della ricerca sono compositi, e vengono elencati in modo omogeneo sulla destra, fruibili in diverse modalità: elenco, galleria, *timeline*, ma anche su mappa, in modo da visualizzare la distribuzione sul territorio delle risorse.

Sulla sinistra, è possibile operare sulle faccette, filtrando per ambito culturale di interesse (che permette di visualizzare soltanto i risultati relativi a uno specifico settore), per localizzazione o per cronologia.

Tra i risultati evidenziati nel filtro per ambito culturale, vediamo anche la scheda di autorità dell'architetto Elia Fornoni.

Gli *authority data* in Beweb

Le schede di autorità all'interno di Beweb fungono da punto di snodo per la navigazione *cross-domain* di banche dati che descrivono risorse diverse. Per la gestione dei punti di accesso è stato adottato il seguente modello di lavoro: i beni vengono descritti dai professionisti presenti sul territorio (bibliotecari, archivisti, architetti, storici dell'arte) rispettando ognuno gli standard propri di settore.

La scelta è stata quella di non adottare uno standard descrittivo comune proprio per garantire il rispetto delle specificità di ciascun ambito.

Le entità persona, ente e famiglia vengono rilevate nell'ambito dei diversi progetti di censimento e catalogazione e si potrebbe configurare dunque il caso che una stessa entità 'persona' (come ad esempio l'architetto Elia Fornoni) sia rappresentata in diversi ambiti catalografici: la soluzione adottata è stata quindi quella di gestire centralmente tutti questi *authority file* di diversa provenienza attraverso un sistema di *clustering* che guida la selezione dei nomi che si riferiscono alla medesima entità.

Quindi anche nel caso dei punti d'accesso la scelta è stata quella di non applicare una sintassi comune per la forma del nome, ma di adottare il modello del *cluster* che troviamo presente ad esempio in VIAF. Si forma in questo modo un grappolo di termini di accesso equivalenti che convergono in un punto d'accesso aggregante (AF CEI), portandosi dietro dati minimi e fonti.

Il punto di accesso AF CEI *cross-domain* è composto da informazioni biografiche e storiche (che compongono il set di dati minimo e obbligatorio per le descrizioni di punti di accesso) e da link di approfondimento (come Wikipedia o lo stesso VIAF). La scheda viene poi integrata con una forma 'amichevole' e diretta del nome.

Ma la parte più interessante della scheda riguarda la possibilità di una navigazione multilivello con i contenuti del portale:

- il collegamento ai beni che intrattengono la stessa tipologia di relazione con l'entità, che consente una navigazione trasversale con tutte le risorse con le quali l'entità ha la stessa relazione (ad esempio autoriale o di soggetto);
- il collegamento ai beni sui quali è possibile un'associazione per prossimità (esplicitata con l'etichetta 'potrebbe interessarti');
- e il link diretto alle basi dati di ambito (esplicitata con l'etichetta 'lo trovi negli ambiti').

Il controllo dei punti di accesso all'interno dell'archivio di autorità consente di avere a disposizione delle informazioni che fungono da punto di partenza per avviare una navigazione dinamica, conducendo ad esempio un appassionato di arte a entrare in contatto con la documentazione archivistica relazionata alla pala d'altare di suo interesse, scoprendo ad esempio che di quella pala d'altare se ne parla in una visita pastorale contenuta in un fondo archivistico di una curia vescovile. Continuando la navigazione si può scoprire la biografia del vescovo che ha condotto quella visita e venire a conoscenza della bibliografia su di lui, oltre a scoprire che lo stesso vescovo ordinò il restauro di una determinata cattedrale di cui si possono sfogliare le immagini.

Tra le soluzioni individuate per agevolare la lettura trasversale del patrimonio culturale ecclesiastico da parte di un pubblico più ampio, non di soli specialisti, c'è quella della semplificazione.

Il linguaggio specialistico degli inventari (rispondente a standard e vocabolari di settore) è stato 'tradotto' mediante una visualizzazione amichevole delle informazioni, veicolate da etichette comprensibili rese con linguaggio naturale.

Beweb è raggiungibile anche da SBN. Se il bene librario cercato è localizzato in una biblioteca del Polo biblioteche ecclesiastiche (PBE), il link al catalogo porta direttamente in Beweb. Dalla scheda libro, alla scheda autore collegata, dove scoprire la presenza di beni di ambito diverso legati alla ricerca fatta.

Le prospettive già in studio per la gestione degli AF CEI sono: l'incremento di nuove categorie di entità (a persone, enti e famiglie verranno aggiunti progressivamente luoghi, termini topic e opere); l'incremento di nuove relazioni (non più solo tra entità e beni, ma tra entità ed entità, come ad esempio le relazioni 'è allievo di', 'è collaboratore di', 'è committente di', 'è successore di', che legano due entità persona, o la relazione 'è vescovo di' che lega una entità persona con una entità ente, oppure ancora la relazione 'è finanziato da' che lega un ente ad una entità famiglia).

ABSTRACT

Gruppo di lavoro sugli *authority file* dell'Ufficio per i beni culturali ecclesiastici e l'edilizia di culto della Conferenza episcopale italiana, composto da Adriano Belfiore, Francesca Maria D'Agnelli, Laura Gavazzi, Claudia Guerrieri, Maria Teresa Rizzo, Silvia Tichetti, Paul Gabriele Weston, sito web <http://www.beweb.chiesacattolica.it/it>, e-mail beweb@chiesacattolica.it.

Beweb rende possibile l'esplorazione del patrimonio culturale delle diocesi italiane fatto di beni storico artistici, librari, archivistici, edifici di culto e istituti culturali, attraverso la ricerca *cross-domain*, gli approfondimenti tematici e il costante *authority work* sulle entità (persona, ente, famiglia), a vario titolo relazionate con i beni e con gli altri contenuti informativi del portale.

La sfida principale? Individuare i punti di accesso per navigare trasversalmente banche dati che descrivono beni culturali di diversa natura, e potenziare le relazioni tra gli oggetti rendendoli capaci non solo di interoperare ma anche di 'illustrarsi e introdursi' vicendevolmente, in modo sinottico e integrato.

La Fototeca Zeri in linked open data

di Marilena Daquino e Francesca Tomasi

Il progetto Zeri & LODE

L'Archivio fotografico Federico Zeri¹ dell'Università di Bologna preserva una delle più prestigiose collezioni di fotografie di opere d'arte d'Europa (290.000 esemplari), documentata da una ricca biblioteca di storia dell'arte (46.000 volumi), da una delle più consistenti raccolte di cataloghi d'asta (37.000 cataloghi), nonché dai preziosi appunti del suo creatore (15.000 documenti archivistici).

Nel 2003 la Fondazione Federico Zeri ha iniziato il riordino del fondo e la contemporanea realizzazione di un catalogo digitale delle fotografie, includendovi il ricco repertorio delle opere riprodotte e descritte nella documentazione archivistica. A tale fine sono stati adottati due standard nazionali, proposti dall'Istituto centrale per il catalogo e la documentazione (ICCD), la Scheda F² per la descrizione delle fotografie e la Scheda OA³ per la descrizione dell'opera d'arte, in questo caso soggetto della fotografia e non oggetto di catalogazione a se stante. Data la complessità e la ricchezza informativa offerta da tali normative, sono stati utilizzati i campi descrittivi corrispondenti al livello minimo inventariale, i quali ammontano a circa 120 campi su 300 per la Scheda F e circa 100 su 280 per la Scheda OA.

Nel 2013 l'International Consortium of Photo Archives (PHAROS)⁴ ha proposto ai suoi 14 membri, tra cui la Fondazione Zeri, la partecipazione attiva nella realizzazione di un ambiente dove riversare immagini e metadati relativi alle opere d'arte. L'obiettivo è fornire agli utenti di riferimento, studiosi e critici dell'arte, un luogo in cui poter beneficiare dell'integrazione di informazioni provenienti dalle più autorevoli collezioni fotografiche europee e nordamericane. Per ovviare alle difficoltà derivanti dalla convivenza di differenti standard descrittivi, di un differente livello di analiticità e qualità dei dati, nonché dalle eterogenee tecnologie utilizzate dai singoli partner, i linked open data e l'adozione dello standard CIDOC-CRM⁵ sono stati scelti quale strumento più efficace per raggiungere l'interoperabilità tra collezioni digitali.

Con il progetto Zeri & LODE (figura 1), la Fototeca Zeri è stato il primo archivio fotografico italiano e primo tra i partner di PHAROS ad aver trasformato e pubbli-

MARILENA DAQUINO, Alma mater studiorum – Università di Bologna, Multimedia center CRR-MM, Bologna, e-mail marilena.daquino2@unibo.it.

FRANCESCA TOMASI, Alma mater studiorum – Università di Bologna, Dipartimento di filologia classica e italianistica, Bologna, e-mail francesca.tomasi@unibo.it.

1 Fondazione Federico Zeri: <<http://www.fondazionezeri.unibo.it/>>.

2 Normativa ICCD Scheda F: <<http://www.iccd.beniculturali.it/index.php?it/473/standard-catalografi/Standard/10>>.

3 Normativa ICCD Scheda OA: <<http://www.iccd.beniculturali.it/index.php?it/473/standard-catalografi/Standard/29>>.

4 PHAROS: <<http://pharosartresearch.org/>>.

5 CIDOC-CRM: <<http://www.cidoc-crm.org/cidoc-crm/>>.

cato i propri dati sotto forma di linked open data⁶. La sperimentazione è stata condotta su un set di oltre 30.000 schede catalografiche di fotografie raffiguranti circa 19.000 opere d'arte moderna. Sono stati utilizzati i modelli concettuali riconosciuti dalle comunità di riferimento (CIDOC-CRM) e modelli in uso da comunità affini, come le SPAR ontologies⁷ per arricchire il potenziale espressivo di CIDOC-CRM con informazioni di natura bibliografica, e HiCO ontology⁸ per descrivere la provenienza di attribuzioni e la metodologia sottostante.

Sono state così realizzate due ontologie speculari, rispettivamente F entry ontology⁹ e OA entry ontology¹⁰, capaci di rappresentare la ricchezza informativa prevista dagli standard Scheda F e Scheda OA. Le ontologie riutilizzano i termini dei modelli già citati e li estendono con nuove classi e predicati, prevedendo a supporto della descrizione l'utilizzo di vocabolari controllati e aperti, al fine di integrare le lacune di CIDOC-CRM – in particolar modo per ciò che concerne la descrizione del processo interpretativo sottostante un'attribuzione di autorialità di un'opera d'arte.

Tra gli ambiti coperti nella rappresentazione formale fornita dai due modelli citati si ritrovano: la descrizione fisica dell'oggetto e la sua collocazione (inclusa la descrizione archivistica del fondo), eventi e ruoli coinvolti nel ciclo di vita dell'oggetto culturale – dalla commissione alla collezione, la riproduzione, la stampa, il cambio di proprietà e i trasferimenti –, le relazioni con altri documenti, allegati o citati quali fonti durante il processo di catalogazione, e infine le attribuzioni, motivate da criteri tecnici o fonti bibliografiche.

Catalog entries	OA Entry		F Entry
Ontologies	OA Entry Ontology	CIDOC-crm SPAR HiCO	F Entry Ontology
Mapping Documents	OA Entry to RDF		F Entry to RDF
Data and exemplars	OA Entry RDF/ttl example	Zeri Archive LOD	F Entry RDF/ttl example

Figura 1 – Rappresentazione grafica dei risultati del Progetto Zeri & LODE

Gli archivi fotografici nel *Web of data*: un modello per l'integrazione dei domini?

Il progetto Zeri & LODE si definisce così come un primo – e *in fieri* – tentativo di operare una reale integrazione semantica tra domini culturali (musei, biblioteche e archivi), ambendo a rappresentare nello scenario dei linked data tipologie eterogenee di informazioni.

6 Zeri & LODE project: <<https://w3id.org/zericatalog/>>.

7 SPAR ontologies: <<http://www.sparontologies.net/ontologies/>>.

8 HiCO ontology: <<http://purl.org/emmedi/hico>>.

9 F entry ontology: <<http://www.essepuntato.it/2014/03/fentry/>>.

10 OA entry ontology: <<http://purl.org/emmedi/oaentry>>.

Tramite la definizione di modelli il più possibile esaustivi e trasversali nel vasto dominio dei beni culturali, si vogliono fornire alcune basi concettuali e tecniche utili a più tipologie di interlocutori. Non è quindi un risultato di esclusiva pertinenza degli archivi fotografici, bensì una *mapping* dell'esistente dove le specificità di ogni singolo dominio e di ciascuna tradizione vengono mantenute nel passaggio al nuovo modello del *semantic Web*. Il fine è rendere effettiva la descrizione di scenari complessi tramite la sperimentazione su un caso d'uso significativo, quale quello fornito dall'archivio Zeri. Pertanto, oltre ai due modelli ontologici citati, sono stati realizzati due documenti¹¹ che approfondiscono il *mapping* dei campi delle Schede F/OA a RDF, affiancati da esempi e casi d'uso, affinché altre tipologie di enti possano utilizzare tali strumenti per la conversione delle proprie descrizioni sulla base di predicati ontologici esistenti.

Con la pubblicazione di un primo *dataset* RDF si vogliono poi offrire i presupposti pratici per il riutilizzo dei dati e la futura integrazione a progetti affini. Il *dataset* attualmente consta di 11 milioni di triple per descrivere principalmente fotografie, opere ritratte, artisti, personalità ed enti coinvolti a vario titolo, eventi legati alla storia degli oggetti culturali e relazioni con documenti collegati (monografie, riviste, documentazione archivistica ecc.). Le principali entità individuate sono collegate a record di *authority file* di riferimento (VIAF, Getty ULAN, GeoNames), ad altri *dataset* (Dbpedia, Wikidata) e risorse web (le schede online del catalogo Zeri e Wikipedia). Anche la descrizione degli oggetti culturali, come il supporto materiale, la tecnica artistica e la tipologia di oggetto, è supportata da link a *thesauri* rilevanti per il settore (Getty AAT).

Lo scopo ultimo del progetto è la definizione di nuovi strumenti per lo studio e la ricerca utilizzando le tecnologie del *semantic Web*, contribuendo innanzitutto a disvelare un patrimonio informativo non ancora espresso in tutto il suo potenziale. Una particolare attenzione è stata rivolta alla descrizione degli aspetti soggettivi sottesi alla catalogazione e l'attribuzionismo – uno degli elementi più significativi data la natura del fondo, focalizzata sulla storia dell'arte. L'analisi documentata, la classificazione archivistica, il riferimento all'opinione di eminenti storici dell'arte e l'accuratezza e la ricchezza delle attribuzioni fornite dalla catalogazione, sono tutte caratteristiche fondamentali che richiedono una adeguata riflessione se inserite nel nuovo panorama dei LOD.

Nello scenario in cui informazioni contraddittorie sullo stesso soggetto possono convivere più o meno pacificamente, dove informazioni più o meno autorevoli possono essere fornite da soggetti più o meno competenti e titolati, garantire la qualità e la provenienza delle informazioni fornite dalle istituzioni culturali deve essere oggetto di attenzione, pena l'inconsistenza dei dati al momento dell'integrazione con altre fonti.

Tale problematica è stata oggetto di formalizzazione nel modello HiCO (Historical context ontology), affinché fosse contemplata la possibilità di gestire informazioni contraddittorie nella futura integrazione tra archivi affini. Perciò particolare rilievo è stato dato alla descrizione della metodologia, delle motivazioni, le fonti e i criteri sottesi alla scelta di una qualsiasi attribuzione (autorialità, datazione, ruolo coinvolto nel ciclo di vita della fotografia ecc.). Questi dati sono infatti considerati garanzia di informazione autorevole, in grado cioè di sottolineare la maggiore qualità dei dati dei beni culturali rispetto al *mare magnum* del *Web of data*, e sono quindi deputati a garantire un futuro corretto reperimento delle informazioni da parte dell'utente finale.

¹¹ Cfr.: Mapping FtoRDF, <<https://dx.doi.org/10.6084/m9.figshare.3175273.v1>>; Mapping OAtoRDF, <<https://dx.doi.org/10.6084/m9.figshare.3175057.v1>>.

ABSTRACT

MARILENA DAQUINO, Alma mater studiorum – Università di Bologna, Multimedia center CRR-MM, Bologna, e-mail marilena.daquino2@unibo.it.

FRANCESCA TOMASI, Alma mater studiorum – Università di Bologna, Dipartimento di filologia classica e italianistica, Bologna, e-mail francesca.tomasi@unibo.it.

L'archivio fotografico Federico Zeri dell'Università di Bologna preserva una delle più prestigiose collezioni di fotografie di opere d'arte d'Europa, una ricca biblioteca di storia dell'arte e una consistente raccolta di cataloghi d'asta. Nel solco del progetto Zeri & LODE e del più ampio consorzio internazionale PHAROS, la Fototeca Zeri è stato il primo archivio fotografico italiano a trasformare i propri dati in linked open data, realizzando due ontologie, F entry ontology e OA entry ontology, così da rappresentare la ricchezza informativa prevista dagli standard catalografici Scheda F e Scheda OA.

Coming AUTH: per una bonifica e implementazione dell'authority file di SBN

di Agnese Galeffi, Andrea Marchitelli, Patrizia Martini e Lucia Sardo

Una delle critiche più comunemente rivolte a SBN riguarda la carenza di dati di autorità strutturati, autorevoli e aggiornati. A questo si aggiunga il fatto che il problema della manutenzione dei dati, sia del catalogo che dell'*authority file*, è di difficile soluzione. La manutenzione è difficile da gestire in un ambito cooperativo come vuole essere quello di SBN. Un esempio su tutti: i livelli, che pure nascevano con lo scopo di creare record con livelli di completezza diversi¹, sono diventati, in alcuni casi, quasi un ostacolo alla cooperazione.

La mancanza di dati autorevoli crea difficoltà nelle attività di catalogazione, ma soprattutto crea difficoltà nell'utente finale dei dati (sia questo un bibliotecario che deve fornire servizi di reference, un bibliotecario addetto agli acquisti, l'utente vero e proprio del catalogo alla ricerca di informazioni o di 'oggetti' fisici o digitali).

Parliamo di dati e record di autorità ma sappiamo che la maggioranza dei punti di accesso in SBN sono record di autorità costituiti dalla forma preferita del nome e dalle forme di rinvio, ove presenti. Non per questo la correzione delle vedette è agevole dal momento che, per modificarle, occorre essere all'interno di un Polo che possieda una delle notizie collegate.

A titolo di esempio, riportiamo i dati relativi al numero di VID, divisi per tipologia e per livello:

	A	B	C	D	E	G	R	Totale
AUF	1.895	592	55.515	2.976	94	6		61.078
SUP	2.882	834	118.416	8.679	7.511	5.153	2.523	145.998
MAX	4.145	1.312	134.237	9.643	24.770	9.282	4.492	187.881
MED	5.812	2.112	438.479	25.932	49.770	15.822	21.542	559.469
MIN	32.136	18.951	2.283.040	144.005	229.864	44.261	82.558	2.834.815
REC	5.256	2016	418.306	29.100	89.449	12.668	20.884	577.679
Totale	52.126	25.817	3.447.993	220.335	401.458	87.192	131.999	4.366.920

Figura 1

AGNESE GALEFFI, Scuola vaticana di biblioteconomia, Città del Vaticano, e-mail galeffi@vatlib.it.

ANDREA MARCHITELLI, EBSCO information services, e-mail amarchitelli@ebSCO.com.

PATRIZIA MARTINI, Istituto centrale per il catalogo unico, Roma, e-mail patrizia.martini@beniculturali.it.

LUCIA SARDO, Alma mater studiorum – Università di Bologna, Vicepresidenza della Scuola di lettere e beni culturali, Ravenna, e-mail lucia.sardo@unibo.it.

¹ Cfr. <http://norme.iccu.sbn.it/index.php/Norme_comuni/Codici/Informazioni_di_servizio/Livelli_di_autorita%C3%A0>.

Dai dati si evince che a fronte di quasi sedici milioni di notizie, ci sono 4.366.920 VID di cui solo 207.076 sono 'record di autorità' con livelli superiori a 90. I record più ricchi di dati sono però carenti in struttura² e disomogenei in stile, linguaggio e lunghezza della nota biografica.

SBN è cresciuto grazie all'avanguardia dell'ottica cooperativa dei suoi esordi ma non ha saputo sfruttare tale modalità di lavoro per la manutenzione dei dati, ritenendola una attività da svolgere pressoché esclusivamente in modo centralizzato. Pensare a un unico *database editor* per una realtà delle dimensioni di SBN era impensabile, quindi l'ICCU ha lanciato nel tempo numerosi progetti di bonifica, già a partire dagli anni Novanta, per arrivare al convegno del 2002³ e al documento *Linee guida per la compilazione dell'authority control - Autore personale in SBN*⁴ presente sul sito ICCU e datato 2015⁵. Ricordiamo anche il documento del 2015 *Sintesi delle priorità emerse dai documenti inviati dai Poli SBN in vista dell'assemblea*⁶.

La genesi del progetto Coming AUTH è recente e si localizza durante il convegno per il trentennale di SBN organizzato dall'ICCU e tenutosi alla BNCR di Roma il 1 aprile 2016, "1986-2016. Trent'anni di biblioteche in rete". In tale occasione è risultato evidente che l'esperienza di SBN era legata ad una generazione alla quale, per motivi non dipendenti dalla loro volontà, non se ne sono potute affiancare altre più giovani. In particolare, di più giovani. Questo potrebbe implicare che quando le cosiddette 'nuove generazioni' si troveranno a gestire SBN (cosa inevitabile, pena la chiusura di SBN o la sua manutenzione da parte di soggetti esterni) non è detto che lo sentiranno come qualcosa di 'loro'.

Coming AUTH propone, nel contesto di SBN, una possibile soluzione, con due modalità di *authority work*: una professionale e una *social*.

Da anni, come accennato sopra, esistono progetti cooperativi in questo settore in SBN (uno dei quali attualmente in corso), la novità di Coming AUTH risiede nel coinvolgimento di professionisti in quanto tali e non in virtù di una loro appartenenza istituzionale.

La modalità professionale vede la collaborazione fra ICCU e AIB Lazio per addestrare all'uso degli specifici strumenti di lavoro (interfaccia diretta, software catalografici come SBN Web) operatori specializzati in grado di bonificare i dati di autorità.

2 Ad esempio, gli estremi cronologici sono riportati in un unico campo, non in un campo data di nascita e in uno data di morte.

3 <http://www.iccu.sbn.it/opencms/opencms/it/main/attivita/naz/pagina_96.html>.

4 <http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/AF_Indicazioni_e_regole_riv_23_luglio2015.pdf>.

5 Trattandosi di attività centralizzate e particolarmente delicate, in quanto incidono direttamente sul catalogo collettivo e su record di altri Poli, l'ICCU per ora ha abilitato tali nuove funzionalità solo ai bibliotecari membri della struttura di manutenzione del catalogo SBN. Si auspica tuttavia la possibilità di allargare il *pool* di lavoro a nuovi Poli, secondo le indicazioni che verranno date dal Comitato tecnico scientifico di SBN, <http://www.iccu.sbn.it/opencms/opencms/it/main/sbn/catalog_manutenz_cat_sbn/pagina_0001.html>.

6 <http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/Assembleapoli_sintesipriorita.pdf>. Cfr. la pagina dedicata ai documenti di proposta presentati dai singoli Poli: <http://www.iccu.sbn.it/opencms/opencms/it/main/sbn/organi_sbn/pagina_0001.html>. Cfr. anche le proposte ICCU: <http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/proposteICCU_assemblea_poli.pdf>.

La scelta delle entità da bonificare prioritariamente verrà effettuata secondo criteri di selezione decisi in base alle esigenze dei diversi Poli coinvolti, in un'ottica di cooperazione a livello regionale (potremmo dire di sussidiarietà). Una prima ipotesi ha fatto emergere la possibilità di intervenire sui VID relativi ai bibliotecari, in modo da proporre una scelta che vada al di là delle specificità dei singoli poli e che sia facilmente identificabile, quantificabile e che consenta di testare la validità delle premesse teoriche. Grazie a repertori esistenti, inoltre, il lavoro sarebbe decisamente semplificato⁷.

Una preliminare estrazione dei dati permetterà di verificare la validità dell'ipotesi e di definire i record da lavorare e i tempi di lavoro. Queste variabili rendono anche il progetto scalabile e affrontabile a tappe, qualora necessario.

La seconda modalità, quella *social*, avverrebbe tramite un servizio *Mix and match* come quello sviluppato da Wikimedia⁸, invitando gli utenti, previa registrazione, a confermare la corrispondenza o meno di entità di diversa natura (nomi, luoghi ecc.) tratte da fonti di diversa natura, non solo bibliotecaria. Questi sistemi fanno leva, in genere, sulla 'gratificazione' derivante dall'attribuzione di *benefit* o *bagde*. Ovviamente sarà prevista una soglia di affidabilità di queste attribuzioni, ad esempio, validandole solo se una determinata percentuale si è espressa a favore e nessuno contro.

A titolo d'esempio, in ambito archivistico Daniel Pitti, in occasione del convegno IFLA del 2016, ha presentato un progetto denominato SNAC (Social Networks and Archival Context), che usa i dati archivistici che descrivono le persone per trovare e identificare le risorse storiche distribuite e costituire così la base per una cooperazione internazionale in questo ambito⁹. A oggi, quasi 25 milioni di entità SNAC trovano corrispondenza nel VIAF permettendo di arricchire le descrizioni SNAC con nomi alternativi e collegamenti a record bibliografici Worldcat.

Nel *Mix and match* di *Coming AUTH social*, una entità sarebbe un record di autorità povero di dati e l'altra da confrontare potrebbe essere o un altro record di SBN (per la deduplicazione) o altre fonti bibliotecarie (VIAF, ISNI ecc.) o meno (Wikipedia, Wikidata).

In Wikipedia gran parte delle voci che riguardano persone contengono, mostrati in un *box* al fondo della voce, codici e link ai record di cataloghi e *authority file* di diversa natura¹⁰.

Il sistema che rende possibile questi collegamenti è basato su proprietà di Wikidata: un apposito template¹¹ aggiunge i dati di *authority* alle voci.

Un progetto di questo tipo potrebbe creare un circolo virtuoso poiché, arricchendo l'*authority file* di SBN, si arricchiscono tutti gli altri progetti, sia Wiki che bibliotecari e, di conseguenza, SBN può utilizzare questi altri dati per arricchire i propri record di autorità. Ad esempio, attraverso il collegamento a Wikidata si potrebbero immettere i dati di altri *authority file* (per esempio, i codici ORCID) in SBN.

La modalità *social* avrebbe anche l'effetto di aprire SBN agli utenti, favorendo la partecipazione di non addetti ai lavori, aumentando la visibilità di SBN e dei progetti a esso collegati, con la possibilità di un coinvolgimento attivo anche del mondo Wikimedia.

7 Si veda, ad esempio, la versione online del *Dizionario bio-bibliografico dei bibliotecari italiani del XX secolo*, a cura di Simonetta Buttò: <<http://www.aib.it/aib/editoria/dbbizo/dbbizo.htm>>.

8 <<https://tools.wmflabs.org/wikidatagame/distributedhttps://tools.wmflabs.org/mix-n-match/>>.

9 Per avere una idea si può consultare la pagina di ricerca del prototipo: <<http://socialarchive.iath.virginia.edu/snac/search>>.

10 Cfr. <https://it.wikipedia.org/wiki/Aiuto:Controllo_di_autorit%C3%A0>.

11 Cfr. <https://it.wikipedia.org/wiki/Template:Controllo_di_autorit%C3%A0>.

Coming AUTH quindi, oltre a migliorare l'*authority file* di SBN, dovrebbe apportare i seguenti vantaggi:

- per la modalità professionale: il progetto è scalabile e affrontabile a tappe, basta scegliere i *dataset* da lavorare e assicurare un ricambio nei gruppi di lavoro;
- per la modalità *social*: coinvolgimento e visibilità di SBN, dell'ICCU e dei suoi progetti, favorendo la partecipazione di non addetti ai lavori, con la possibilità di un coinvolgimento attivo anche del mondo Wikimedia. La promozione di SBN in ambiti non bibliotecari è un vantaggio non trascurabile e va incontro a esigenze di promozione e marketing a costo zero.

L'innovazione di Coming AUTH sta nell'affrontare un problema vecchio usando soluzioni nuove.

ABSTRACT

AGNESE GALEFFI, Scuola vaticana di biblioteconomia, Città del Vaticano, e-mail galeffi@vatlib.it.

ANDREA MARCHITELLI, EBSCO information services, e-mail amarchitelli@ebSCO.com.

PATRIZIA MARTINI, Istituto centrale per il catalogo unico, Roma, e-mail patrizia.martini@beniculturali.it.

LUCIA SARDO, Alma mater studiorum – Università di Bologna, Vicepresidenza della Scuola di lettere e beni culturali, Ravenna, e-mail lucia.sardo@unibo.it.

Il progetto Coming AUTH è un progetto volto all'implementazione dell'*authority file* di SBN, con una modalità professionale e una modalità aperta al mondo non bibliotecario. La modalità professionale, grazie alla collaborazione fra ICCU e AIB Lazio, prevede il coinvolgimento di giovani professionisti. La scelta delle entità da bonificare verrà decisa in base alle esigenze dei diversi poli coinvolti, in un'ottica di cooperazione a livello regionale.

La modalità *social*, avverrebbe tramite un servizio *Mix' n' match* come quello sviluppato da Wikimedia, invitando gli utenti, previa registrazione, a confermare la corrispondenza o meno di entità di diversa natura tratte da fonti diverse. La modalità *social* vorrebbe anche aprire SBN agli utenti, favorendo la partecipazione di non addetti ai lavori, aumentando la visibilità di SBN e dei progetti a esso collegati, con la possibilità di un coinvolgimento attivo anche del mondo Wikimedia.

Articolo proposto il 7 dicembre 2016 e accettato il 22 febbraio 2017.

ABSTRACT

AIB studi, vol. 57 n. 1 (gennaio/aprile 2017), p. 91-117. DOI 10.2426/aibstudi-11569

GRUPPO DI STUDIO AIB CILW (Catalogazione, indicizzazione, linked open data e web semantico), composto da Roberto Raieli, Marinella Cisternino, Elena Corradini, Antonella Iacono, Antonella Trombone, sito web <www.aib.it/struttura/commissioni-e-gruppi/cnc>, e-mail cnc@aib.it. Home page del Convegno AIB CILW 2016: <<http://www.aib.it/attivita/congressi/c2016/giornata-studi-aib-cilw-2016/>>.

L'universo delle risorse culturali: lampi di genio e azioni concrete. *Lightning talks* presentati al Convegno AIB CILW 2016

Sono raccolti sei dei sette *lightning talks* presentati al Convegno AIB CILW 2016, nell'ambito della sessione aperta al contributo di qualunque studioso delle tematiche proposte nell'incontro. Le relazioni hanno seguito lo stile del *lightning talk*, brevi e incisivi commenti di tipo teorico o dimostrazioni di carattere pratico. Gli speaker hanno dato le proprie risposte ai principali interrogativi del convegno, quali: nello spazio del Web dei dati, che tipo di convergenza è realmente possibile attuare tra i diversi istituti culturali che si occupano della gestione delle risorse della conoscenza? Ci sono realmente i presupposti storici, teorici e pratici per ripensare in tal senso le strategie dell'informazione? Qual è lo stato dell'arte dei progetti MAB (Musei, archivi e biblioteche) e GLAM (Galleries, libraries, archives, museums)?

The universe of cultural resources: between eureka and concrete actions. *Lightning talks* presented to AIB CILW 2016 Conference

Here are collected six out of the seven lightning talks presented at the AIB CILW 2016 Conference, as part of the session open to the contribution of any scholar dealing with the topics of the meeting. The presentations followed the lightning talk style, brief and incisive theoretical comments or practical demonstrations.

The speakers gave their answers to the main questions of the conference, such as: inside the Web of data, what kind of convergence is to be implemented among the different cultural institutions that are in charge of the organization and curation of knowledge resources? Are there sufficiently convergent historical, theoretical, and practical bases to build on and rethink about information strategies? Which is the state-of-the-art of the GLAM projects (Galleries, libraries, archives, museums)?