

MATERIALI

La convergenza delle risorse informative nel Web dei dati e delle relazioni (note sul Convegno AIB CILW 2016, Biblioteca nazionale centrale di Roma, 21 ottobre 2016)

di Luigi Catalani

Introduzione

I recenti, importanti sviluppi delle tecnologie per il trattamento dell'informazione hanno sollecitato gli istituti culturali che si occupano della gestione delle risorse del sapere a escogitare nuove modalità di cooperazione, capaci di far fronte alla crescente mole di informazione prodotta e alla diminuzione degli investimenti finanziari nel settore, che negli ultimi anni stanno condizionando la gestione e la cura delle risorse della conoscenza. Il fenomeno più generale della convergenza al digitale si traduce oggi, per questi istituti appartenenti a diversi livelli istituzionali, nella condivisione di una piattaforma tecnologica in grado di rispondere in maniera flessibile all'esigenza comune del trattamento delle risorse informative. L'adozione generalizzata del Web semantico e dei linked open data (LOD) sembra, in effetti, poter mettere nelle condizioni gli operatori di biblioteche, musei e archivi di offrire un servizio più efficace e puntuale per l'accrescimento culturale della società mediante l'accesso e la scoperta di un'ampia gamma di risorse della conoscenza. Lo sviluppo di una serie di progetti di gestione condivisa delle risorse, basati sulla reale integrazione tra i dati di diversa tipologia e provenienza, testimonia la fiducia delle istituzioni e degli addetti ai lavori nel Web dei dati come linguaggio condiviso per la produzione e la gestione coordinata delle risorse della conoscenza.

Un'occasione per riflettere sui reali presupposti storici, teorici e pratici di questo ripensamento delle strategie del mondo dell'informazione è stata offerta dal "Convegno AIB CILW 2016" intitolato "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 e patrocinato dall'Associazione italiana biblioteche, Sezione Lazio e dalla Biblioteca

LUIGI CATALANI, Biblioteca provinciale, Potenza; Università degli studi di Salerno, Dipartimento di scienze del patrimonio culturale, Fisciano, e-mail lcatalani@unisa.it.

Ringrazio Antonella Trombone per i preziosi suggerimenti e le sue note di lettura.

Ultima consultazione siti web: 26 febbraio 2017.



nazionale centrale di Roma, che lo ha ospitato il 21 ottobre dello scorso anno¹. Un nutrito gruppo di studiosi e professionisti dell'ambito della Library and information science (LIS) e delle discipline collegate ha provato a definire le opportunità, i requisiti e le criticità della convergenza delle istituzioni culturali sul terreno del Web dei dati. L'obiettivo principale degli organizzatori – cui va il plauso per l'organizzazione di un evento di alto livello professionale e la gestione di una mole notevole di interventi concentrati in una sola giornata – era quello di riunire esperienze, ricerche e studi sul tema dell'innovazione teorica, metodologica, tecnologica e professionale in area MAB (Musei archivi biblioteche) e LAMMS (Libraries, archives, museums, monuments and sites), con particolare riguardo alle tecnologie dei linked data e al progetto del Web semantico.

Teorie e prassi dell'interoperabilità semantica

Dopo i saluti istituzionali, la prima parte del convegno, coordinata da Alberto Petrucciani (Sapienza Università di Roma), ha proposto una serie di riflessioni teoriche, organiche e innovative, utili a chiarire i tratti fondamentali dell'evoluzione dagli OPAC alla navigazione semantica e gli aspetti principali del tema della convergenza delle risorse informative tramite i linked data.

A partire dall'esame degli spazi della memoria, ossia degli ambienti di mediazione informativa realizzati in età moderna dalle istituzioni e nelle istituzioni della conoscenza registrata, Maurizio Vivarelli (Università degli studi di Torino) ha messo in luce alcuni degli aspetti connessi alla dimensione visiva dei dati e degli oggetti documentari ovvero della loro rappresentazione. Se nei secoli XVI e XVII l'organizzazione della conoscenza è esteriorizzata e resa visibile nei cosiddetti teatri della memoria, spazi insieme fisici e concettuali, la tendenza attuale è l'eccentricità, in quanto la visualizzazione delle informazioni (il riferimento è naturalmente ai modelli di visualizzazione di triple RDF) prescinde dallo spazio fisico dei documenti e si materializza altrove. Al contempo, l'architettura di alcune biblioteche contemporanee riflette il disordine, la fluidità e l'ibridazione dell'infosfera nell'era postmoderna. Ogni operazione di mappatura cognitiva e visiva cerca di soddisfare l'esigenza schiettamente umana di fare 'mente locale', ossia di costruire una rappresentazione dei dati che aiuti a orientarsi nell'ecosistema informativo. La mediazione documentaria, che gli utenti sperimentano servendosi delle interfacce e dei metadati predisposti da bibliotecari e archivisti, si colloca dunque in uno spazio che oscilla tra la dimensione interna della mente e la dimensione esterna degli oggetti documentari. Di qui l'invito, sicuramente condivisibile, a favorire l'incontro tra culture documentarie e culture comunicative, a valorizzare la dimensione visiva dei sistemi documentari, a progettare sistemi documentari 'olistici', a reimmaginare la mediazione elaborando schemi, interfacce e presentazioni grafiche tali da facilitare la percezione dell'organizzazione concettuale dei dati e delle informazioni e favorirne l'elabo-

¹ Il comitato scientifico era composto da Enzo Borsellino (Università Roma Tre), Simonetta Buttò (Istituto centrale per il catalogo unico), Silvana de Capua (Biblioteca nazionale centrale di Roma), Andrea De Pasquale (Biblioteca nazionale centrale di Roma), Maria Guercio (Sapienza Università di Roma), Andrea Marchitelli (CINECA), Alberto Petrucciani (Sapienza Università di Roma), Vittorio Ponzani (Istituto superiore di sanità), Giovanni Solimine (Sapienza Università di Roma) e da Marinella Cisternino (Istituto Regina Elena), Elena Corradini (Biblioteca comunale di Ala), Antonella Iacono (Biblioteca civica di Biella), Roberto Morelato (Cooperativa Libriliberi Bolzano), Roberto Raieli (Università Roma Tre) e Antonella Trombone (Università della Basilicata).

razione in un contesto che può essere anche completamente autonomo rispetto a quello originario².

Paul Gabriele Weston (Università degli studi di Pavia) si è soffermato sul processo di *reductio ad unum* che vede al centro gli *authority data*, elementi importanti già nell'epoca analogica, diventati fondamentali nell'universo digitale per il reperimento dell'informazione e la gestione delle ambiguità. Nel Web dei dati e delle relazioni confluiranno masse sempre più grandi di dati provenienti dai cataloghi delle biblioteche e di altre istituzioni culturali ma sganciati dagli archivi di provenienza³. A partire dal riconoscimento del ruolo decisivo che gli identificativi rivestono nel contesto dei linked open data, per la loro capacità di individuare una precisa entità (o un *cluster* di identificativi) in maniera univoca e persistente, Weston ha tracciato l'evoluzione dei record di autorità nella pubblicistica scientifica e mostrato la tendenza all'unicità, alla convergenza e alla coesistenza dei formati attraverso l'esame degli *authority file* di una serie di biblioteche digitali, portali ad aggregazione e servizi di *e-reference*. Nei casi migliori, l'utilizzo degli identificativi in sistemi costituiti da insiemi di dati eterogenei consente la conservazione del contesto di provenienza del singolo dato e favorisce la *cross-fertilisation*, ossia la contaminazione di idee e concetti provenienti da ambiti diversi capace di generare beneficio reciproco e nuova conoscenza.

Maria Teresa Biagetti (Sapienza Università di Roma) ha fornito un quadro delle opportunità che le tecnologie del Web semantico concedono per la realizzazione di strutture per lo scambio e l'integrazione delle informazioni appartenenti a fonti eterogenee del patrimonio culturale. Le ontologie sono strumenti preziosi per facilitare l'accesso alle informazioni e agli oggetti digitali gestiti con modalità diverse dai musei, dagli archivi o dalle biblioteche. Accanto a modelli concettuali generali come CIDOC-CRM (International Committee for Documentation – Conceptual Reference Model)⁴, un'ontologia formale che si è imposta come modello concettuale a livello internazionale, si vanno affermando altri modelli per l'interoperabilità semantica, come gli OBDA (Ontology based data access). Nel Web dei dati le ontologie esprimono le relazioni semanticamente qualificate su cui si fonda la relazione tra i dati. Nell'ambito del patrimonio culturale, CIDOC-CRM è particolarmente utile per la realizzazione di *interlinking* tra entità e proprietà. Nella stessa direzione vanno i progetti coordinati dall'ICCU (Athena, Linked heritage, Athena plus) volti a facilitare l'accesso integrato alle collezioni digitali e la partecipazione a Europeaana.

Con elegante ironia, Paola Castellucci (Sapienza Università di Roma) ha segnalato l'esigenza di impegnarsi per una disciplina etica dell'accesso all'informazione e di fermare il dilagare del 'sillogismo di Amazon', in base al quale feedback, commenti e consigli per gli acquisti indirizzano subdolamente le nostre ricerche. In base agli algoritmi

2 Cfr. *Data visualization: a guide to visual storytelling for libraries*, edited by Lauren Magnuson. London: Rowman & Littlefield, 2016; Manuel Lima, *The book of trees: visualizing branches of knowledge*. New York: Princeton Architectural Press, 2014.

3 Cfr. *Global interoperability and linked data in libraries: special issue*, «JLIS», 4 (2013), n. 1; Simona Turbanti, *Cui prodest libraries authority work?*, «JLIS», 5 (2014), n. 2, p. 49-59, DOI: 10.4403/jlis.it-10062; Sabina Cavicchi, *Authority control: aspetti operativi in un contesto universitario e nuove esperienze*, «JLIS», 6 (2015), n. 1, p. 161-182, DOI: 10.4403/jlis.it-10340.

4 CIDOC-CRM: <<http://www.cidoc-crm.org/>>. Cfr. Martin Doerr; Dolores Iorizzo, *The dream of a global knowledge network: a new approach*, «Journal on computing and cultural heritage», 1 (2008), n. 1, p. 1-23, DOI: 10.1145/1367080.1367085.

prodotti nell'ambito della Ant colonies optimization (ACO) (un sottoinsieme dell'intelligenza artificiale, in particolare della teoria degli sciami), *cluster*, *ranking* e *like* alimentano il *loop* informativo di una comunità di utenti sempre più simili a formiche elettriche, prede dei propri 'ferormoni digitali' rilasciati lungo il percorso, dell'omologazione e della massificazione dei gusti, del monitoraggio delle azioni a fini commerciali, del canone delle vendite come criterio unico. C'è ancora spazio, ha ricordato Castellucci rifacendosi all'insegnamento di Norbert Wiener⁵, per un pensiero divergente, una connessione ludica, un modello stocastico, un individuo libero, capace di rinunciare ai molteplici condizionamenti e di fuoriuscire dal determinismo informativo.

I linked data sono sempre più al centro dei progetti delle più importanti biblioteche europee (Bibliothèque nationale de France, British Library, Deutsche Nationalbibliothek, Biblioteca nacional de España), che sulla base comune di FRBR hanno elaborato diverse interfacce e modalità di rappresentazione dei dati, nell'ottica di facilitare la ricerca dell'utente e di arricchire i *dataset* tramite i collegamenti ad altri *set* di informazioni⁶. Secondo Mauro Guerrini e Carlotta Vivacqua (Università di Firenze), la scomparsa del termine 'catalogo' dal portale aggregatore di dati della BnF⁷ – che fornisce informazioni relative ad autori, opere e soggetti – non preannuncia la morte del catalogo per mano della tecnologia linked open data bensì la sua integrazione in un servizio informativo più vasto e strutturato e soprattutto più adatto al contesto del Web attuale. La BnF provvede inoltre a una diffusione capillare dei dati tramite apposite tecnologie, rende i dati disponibili al riuso, prevede l'adozione di varie ontologie, favorisce i collegamenti a piattaforme e progetti esterni (tra cui WorldCat, VIAF, Wikipedia), detta le fasi di trasformazione dei dati originali in RDF e la loro pubblicazione in pagine HTML facilmente ricercabili dai motori di ricerca. Uno dei frutti di questo lavoro è OpenCat, un catalogo navigabile che utilizza gli identificatori univoci BnF, arricchiti dai collegamenti a risorse esterne⁸. Il progetto omologo della Biblioteca nacional de España⁹, che rispecchia quello della BnF, ha reso disponibili i link alle copie digitalizzate della Biblioteca digital hispánica¹⁰. Il progetto della British Library pubblica una parte della *British national bibliography* in linked data, ponendo al centro del modello la risorsa e collegandola a diversi tipi di entità e di elementi¹¹. La prospettiva è quella dell'arricchimento dei *dataset* tramite ulteriori collegamenti esterni, compresi quelli con altre bibliografie nazionali, come quella della Deutsche Nationalbibliothek, la quale a sua volta, nel giro di pochi anni, ha prima convertito gli *authority file* in linked open data e poi pubblicato in RDF i dati delle sue principali collezioni di monografie e seriali¹².

5 Cfr. Norbert Wiener, *Introduzione alla cibernetica: l'uso umano degli esseri umani*. Torino: Bollati Boringhieri, 2012.

6 Cfr. Mauro Guerrini; Tiziana Possemato, *Linked data per biblioteche, archivi e musei: perché l'informazione sia del web e non solo nel web*. Milano: Editrice Bibliografica, 2015.

7 [Data.bnf.fr: <http://data.bnf.fr/>](http://data.bnf.fr/).

8 http://www.bnf.fr/fr/professionnels/web_donnees_applications_bnf/a.opencat.html.

9 [Datos.bne.es: <http://datos.bne.es/inicio.html>](http://datos.bne.es/inicio.html).

10 <http://www.bne.es/es/Catalogos/BibliotecaDigitalHispanica/Inicio/index.html>.

11 The British National Bibliography as linked open data: <http://bnb.data.bl.uk/>.

12 Linked data service of the German National Library: http://www.dnb.de/EN/Service/DigitaleDienst/LinkedData/linkedata_node.html.

Con un apprezzabile approccio pragmatico, Klaus Kempf (Bayerische Staatsbibliothek) ha posto l'attenzione sull'investimento iniziale necessario per avviare un progetto serio centrato su linked data e semantic Web, a partire dalla creazione di metadati standard e di alta qualità e dalla produzione di *authority file* condivisi dalle principali istituzioni culturali. La qualità dei metadati deriva dal codice di catalogazione (RDA)¹³, dal formato dei dati (MARC 21) e dai record di autorità, che devono avere identificatori persistenti per poter produrre valore aggiunto in termini di facilitazione dei processi di catalogazione, standardizzazione e collegamento. Kempf ha fatto riferimento in particolare al Gemeinsame Normdatei (GND), un sistema di controllo di autorità gestito dalla Deutsche Nationalbibliothek in collaborazione con altre biblioteche di lingua tedesca, per la catalogazione di dati (relativi a persone, persone giuridiche, enti, opere, argomenti, congressi ed eventi, elementi geografici) che sempre più spesso vengono utilizzati anche da archivi, musei e applicazioni Web¹⁴. Tra i progetti di cooperazione resi possibili dall'utilizzo del GND, Kempf ha citato la Deutsche Biographie¹⁵, che grazie all'utilizzo di BEACON (un semplice formato di file che consente di arricchire il patrimonio informativo mediante l'inserimento di link a risorse web) si è evoluta in un articolato sistema informativo storico-biografico, la Deutsche digitale Bibliothek¹⁶, che ha attivato il servizio '*entity facts*' per arricchire le pagine biografiche con i dati del GND, Bavarikon¹⁷ (il portale della cultura e dell'arte bavaresi) e Wikipedia, che nella sua versione tedesca¹⁸ ha attivato la funzione 'cerca persone' (Personensuche)¹⁹. Tutto questo ha un costo, ha concluso Kempf, un costo di gestione che è destinato ad aumentare, in un periodo nel quale le risorse a disposizione delle biblioteche diminuiscono: forse solo lo sviluppo di appositi metodi di apprendimento automatico (*machine learning*) potrebbe ridurre questa forbice che va allargandosi sempre di più.

Buone pratiche di condivisione delle risorse culturali digitali

Seppur concentrata in meno di un'ora, la seconda parte del convegno, coordinata da Andrea Marchitelli (CINECA) e organizzata tramite una *call for proposals* in una serie di *lightning talks* (un esperimento sicuramente riuscito), ha offerto notevoli spunti a partire dalla condivisione di idee e dalla presentazione di esperienze, anche internazionali. Tra queste va annoverata quella descritta da Ruben Almarza e Ricardo Eito-Brun (Universidad Carlos III de Madrid), che hanno esposto i criteri per la selezione di un *repository* semantico per la gestione di dati in un linguaggio SKOS (Simple knowledge organisation system), in grado di sviluppare specifiche e standard per supportare l'uso dei sistemi di organizzazione della conoscenza nel Web

13 Cfr. Carlo Bianchini; Mauro Guerrini, *Introduzione a RDA: linee guida per rappresentare e scoprire le risorse*. Milano: Editrice Bibliografica, 2014. Cfr. Mauro Guerrini; Carlo Bianchini, *Manuale RDA: lo standard di metadatozione per l'era digitale*, in appendice AACR2 vs RDA, di Tiziana Possemato. Milano: Editrice Bibliografica, 2016.

14 Gemeinsame Normdatei (GND): <http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html>.

15 Deutsche Biographie: <<https://www.deutsche-biographie.de/>>.

16 Deutsche digitale Bibliothek: <<https://www.deutsche-digitale-bibliothek.de>>.

17 Bavarikon: <<http://www.bavarikon.de/>>.

18 Wikipedia – Die freie Enzyklopädie: <<https://de.wikipedia.org/>>.

19 Personensuche: <<https://tools.wmflabs.org/persondata/index.php>>.

semantico. L'adozione del linguaggio SKOS presuppone la conversione dei dati in formato RDF e la loro apertura online, in modo tale che possano essere utilizzati da altre applicazioni web. È la ragione per cui oggi molte agenzie sono impegnate nell'opera di conversione di tesauri e schemi di classificazione in linguaggio SKOS e di pubblicazione dei dati in un formato (l'HTML) che però non ne consente il riutilizzo da parte delle macchine. Dall'esame delle funzioni di inserimento, ricerca ed estrazione dei dati emergono diversi aspetti problematici, quali la mancanza di integrazione con gli strumenti per la modifica di dati RDF/SKOS, la complessità degli *end-point* SPARQL e la difficoltà nel gestire dati che non sono a base RDF. Sulla base della valutazione di due strumenti per l'archiviazione semantica (GraphDB²⁰ e Virtuoso²¹), la scelta cade sui database XML. L'integrazione di SKOS e SRU (Search/Retrieve via URL) in un ambiente collaborativo distribuito per la descrizione nello standard EAD (Encoded archival description) garantisce, infatti, maggiore visibilità ai tesauri e ai sistemi di classificazione e stimola tanto lo sviluppo di linguaggi aperti di indicizzazione quanto l'integrazione e il riuso di strumenti per la creazione dei metadati.

Un'altra esperienza internazionale è quella riportata da Filipe Ferreira, Rodolfo Matos, Susana Medina, Maria Manuela Pinto, Augusto Ribeiro e João Rua (Universidade do Porto. Museu digital), che hanno illustrato il progetto U. Porto digital museum²². Basato su una piattaforma tecnologica dinamica e scalabile, U.OpenLab, il progetto supporta il processo di co-creazione e di condivisione dei contenuti digitali, collegati con gli oggetti e i metadati delle collezioni dell'Università di Porto, che intende dunque promuovere il proprio patrimonio culturale e scientifico attraverso l'aggregazione del patrimonio esistente, mediante lo sviluppo di un unico sistema informativo. Il progetto (che si basa sulla normalizzazione attraverso l'utilizzo di standard come il CIDOC-CRM per i metadati e SPECTRUM 4.0 per le procedure museali) promuove un approccio teorico interdisciplinare che accorpa la museologia, le scienze dell'informazione e i sistemi informativi, e prevede inoltre lo sviluppo di una piattaforma tecnologica dinamica e scalabile chiamata OpenLab, che consentirà alla comunità accademica e agli utenti esterni di arricchire le collezioni e di partecipare al processo di co-creazione e (ri)utilizzo dei contenuti digitali dell'Università.

Angela Bellia e Fiammetta Sabba (Università degli studi di Bologna) hanno presentato i metodi e gli strumenti creati nell'ambito del Getty vocabularies as linked open data, un progetto volto a favorire la scoperta, la conoscenza e l'interpretazione del patrimonio di risorse del Getty Research Institute attraverso l'applicazione della tecnologia LOD²³. Il progetto rientra nella strategia di apertura degli archivi digitali del Getty, con l'obiettivo di contribuire a rendere visibile la ricca e diversificata rete di connessioni tra le istituzioni culturali. Grazie alla tecnologia LOD, i Getty vocabularies (*Art & architecture thesaurus*, *Getty thesaurus of geographic names*, *Union list of artist names*, *Cultural objects name authority*) possono essere esplorati mediante relazioni associative e collegamenti impliciti, aumentando l'efficacia dell'indicizzazione e del recupero di informazioni artistiche²⁴. I Getty vocabularies, che condividono una struttura base di dati controllati,

20 GraphDB: <<http://www.graphdb.net/>>.

21 OpenLink Virtuoso: <<https://virtuoso.openlinksw.com/>>.

22 U. Porto Digital Museum: <<https://museudigitalup.wordpress.com/#mdupen>>.

23 Getty vocabularies as linked open data: <<http://www.getty.edu/research/tools/vocabularies/lod/>>.

24 The Getty vocabularies: <<http://vocab.getty.edu/>>.

rappresentano dunque un'interessante pratica per la pubblicazione dei dati aperti della ricerca e l'incremento dello studio collaborativo e della condivisione dei dati²⁵.

Barbara Bonino, Elena Borgi, Maria Pia Girelli, Gabriella Morabito, Federico Morando, Emanuela Secinaro, Luisa Schiavone e Anna Maria Viotto (CoBiS – Coordinamento delle biblioteche speciali e specialistiche dell'area metropolitana torinese) hanno descritto il progetto linked open data messo a punto dal CoBiS²⁶. Frutto della collaborazione di 62 biblioteche, tra cui sette istituzioni pilota che utilizzano quattro diversi software di gestione, il progetto (finanziato dalla Regione Piemonte) utilizza tecnologie *open source* e presenta un sistema di navigazione ispirato al modello di WorldCat, con le integrazioni della LOD cloud che moltiplicano all'infinito le possibilità di esplorazione delle risorse e i collegamenti a collezioni digitali eterogenee già presenti sul Web e ad *authority file* internazionali.

Francesca Maria D'Agnelli, Laura Gavazzi, Adriano Belfiore, Claudia Guerrieri, Maria Teresa Rizzo e Silvia Tichetti (Conferenza episcopale italiana. Ufficio nazionale per i beni culturali ecclesiastici. Gruppo di lavoro sugli *authority file*) hanno mostrato un esempio di utilizzo degli *authority file* per l'integrazione dei beni culturali e una lettura trasversale dei beni culturali della Chiesa cattolica italiana. BeWeb, il portale integrato dei beni culturali ecclesiastici promosso dalla Conferenza episcopale italiana, consente l'accesso a risorse librerie, archivistiche, storiche ed artistiche attraverso una modalità di consultazione *cross-domain*²⁷. La novità consiste nello sviluppo di elementi di contesto, nell'incremento delle relazioni tra gli oggetti, nella conseguente estensione delle potenzialità narrative dei beni²⁸. La circolarità della navigazione è garantita appunto dalle schede di autorità: la notizia AF CEI *cross-domain*, integrata da una visualizzazione amichevole del nome, è corredata da informazioni biografiche e storiche e da collegamenti ad altri *authority file* (come VIAF) o a risorse di approfondimento come Wikipedia²⁹. I futuri sviluppi prevedono l'incremento di nuove categorie ed entità (luoghi, termini tipici, titoli), in aggiunta a quelle già esistenti (persone, enti, famiglie), e di nuove relazioni (non solo tra entità e beni, ma anche tra entità ed entità).

Marilena Daquino e Francesca Tomasi (Università degli studi di Bologna) hanno presentato Zeri & LODE, il progetto linked open data dell'archivio fotografico di Federico Zeri, frutto della collaborazione tra la Fondazione Federico Zeri (che ha aderito al consorzio internazionale degli archivi fotografici PHAROS) e un team di informatici ed esperti in *digital humanities* dell'Università di Bologna³⁰. La conversione

25 Cfr. Patricia Harpring, *Introduction to controlled vocabularies: terminologies for Art, Architecture, and other cultural works*. Los Angeles: Getty Research Institute, 2010, <<http://www.getty.edu/publications/virtuallibrary/160606018X.html>>.

26 CoBiS: <<http://www.cobis.to.it/>>.

27 BeWeB: <<http://www.beweb.chiesacattolica.it/>>.

28 Cfr. Paul Gabriele Weston [et al.], *Gli authority data e l' intersezione cross-domain nei portali ad aggregazione: il portale BeWeb*, «JLIS», 8 (2017), n. 1, p. 138-154, DOI: 10.4403/jlis.it-12127.

29 Cfr. Stefano Russo, *BeWeb: the cross portal of cultural ecclesiastical heritage*, «JLIS», 5 (2014), n. 2, p. 147-157, DOI: 10.4403/jlis.it-10067.

30 Cfr. Ciro Mattia Gonano [et al.], *Zeri e LODE: extracting the Zeri photo archive to linked open data: formalizing the conceptual model*. In: *Proceedings of the 2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL 2014)*. Washington: IEEE, 2014, p. 289-298, DOI: 10.1109/JCDL.2014.6970182. Cfr. Marilena Daquino [et al.], *Enhancing semantic expressivity in the cultural heritage domain: exposing the Zeri photo archive as linked open data*, arXiv.org, 2015, <arXiv:1605.01188> (estensione dell' articolo precedente).

dei dati del catalogo Fototeca Zeri in linked open data intende naturalmente favorire l'accessibilità e il riutilizzo. Per rispondere alle indicazioni del consorzio PHAROS³¹, che promuove la creazione di una piattaforma comune per la ricerca su immagini e metadati relativi alle opere d'arte, il gruppo di lavoro ha scelto i linked open data come veicolo per la condivisione dei dati e l'integrazione con link ad *authority* e *dataset* disponibili in questo formato, e ha mappato i dati sullo standard internazionale CIDOC-CRM, integrandolo con altre ontologie capaci di tradurre la complessità dell'oggetto fotografico e garantire la più ampia interoperabilità semantica³². Un'apposita interfaccia consente l'esplorazione dei primi dati tradotti in triple RDF, che hanno già superato gli 11 milioni.

Agnese Galeffi (Scuola vaticana di biblioteconomia), Andrea Marchitelli (CINECA), Patrizia Martini (ICCU) e Lucia Sardo (Università degli studi di Bologna) hanno comunicato l'inizio di un'operazione di bonifica e implementazione dell'*authority file* di SBN, battezzata Coming AUTH, con una modalità professionale che prevede il coinvolgimento di giovani professionisti sotto la guida dell'ICCU e di AIB Lazio, e una modalità social, più aperta (ancora allo stato progettuale), che intende favorire la partecipazione dei non addetti ai lavori, che saranno invitati ad effettuare un *mix'n-match* delle risorse presenti, allo scopo di deduplicarle e arricchirle. L'effetto di questa duplice attività dovrebbe essere l'incremento della scalabilità e della visibilità di ICCU e SBN e il coinvolgimento attivo anche del mondo Wikimedia.

L'integrazione delle risorse bibliografiche nel Web dei dati

La terza sessione, coordinata da Simonetta Buttò (Istituto centrale per il catalogo unico), è stata orientata alla discussione di progetti ed esperienze di convergenza, capaci di esemplificare le potenzialità applicative dei linked data e del Web dei dati. Laura Crociani ed Elisabetta Viti (Biblioteca nazionale centrale di Firenze) hanno descritto le strategie di mappatura e di integrazione del *Nuovo soggetto*³³ e della WebDewey³⁴. Il primo (composto ad oggi da 60.000 termini) è mappato e navigabile con la DDC per facilitare l'accesso alle collezioni della BNCf, della BNI e dell'Indice SBN, per sfruttare le potenzialità del *Thesaurus* e per rendere i due sistemi reciprocamente navigabili. Il *Nuovo soggetto* colloquia con altri strumenti di organizzazione della conoscenza (tesauri, classificazioni, enciclopedie), e viene sperimentato dal MAB Toscana per l'indicizzazione semantica di documenti archivistici e oggetti museali. A un numero crescente di concetti/termini (più di 10.000) è stata assegnata una notazione Dewey in un campo dedicato, sia in fase di strutturazione dei termini sia in fase retrospettiva. I criteri di mappatura sono lo standard internazionale ISO 25964-2 e le raccomandazioni del Gruppo europeo di utilizzatori della DDC (EDUG)³⁵. I possibili sviluppi vanno nella direzione del potenziamento dell'interoperabilità con possibilità di navigazione reciproca (anche da WebDewey al *Thesaurus*), dell'adeguamento linguistico fra i due strumenti con reciproco arricchimento.

31 PHAROS: the international consortium of photo archives: <<http://pharosartresearch.org/>>.

32 Zeri&LODE: <<http://data.fondazionezeri.unibo.it/>>.

33 Nuovo soggetto – SBN. <<http://thes.bncf.firenze.sbn.it/>>.

34 Versions of the DDC – OCLC: <<http://www.oclc.org/en/dewey/features.html>>. Cfr. Piero Cavaleri, *WebDewey italiana: una nuova edizione in una nuova veste*, «Biblioteche oggi», 32 (2014), n. 1, p. 25-35, DOI: 10.3302/0392-8586-201401-025-1.

35 European DDC Users Group: <<http://edug.pansoft.de/tiki-index.php>>.

chimento lessicale e delle implementazioni di funzioni di reciproca navigabilità con le notazioni di specifiche notizie bibliografiche dell'OPAC BNCf.

Nell'ambito delle biblioteche accademiche, Loredana Cerrone (Università degli studi del Sannio) e Patrizia De Martino (Università degli studi di Salerno) hanno dedicato il loro intervento a SHARE catalogue, un progetto già ben noto tra gli addetti ai lavori, orientato alla cooperazione territoriale e alla condivisione dei servizi bibliotecari e documentari³⁶. Nato nell'ambito di SHARE campus, un sistema integrato di fruizione e gestione delle risorse tra sette università meridionali, è coordinato da Roberto Delle Donne (Università di Napoli Federico II)³⁷. Gli obiettivi principali di SHARE catalogue sono la conversione dei cataloghi bibliografici in linked open data secondo il modello RDF, la pubblicazione su un unico portale dell'intero *dataset*, lo sviluppo di un'unica interfaccia per la consultazione dei dati dei cataloghi bibliografici, l'arricchimento dei dati mediante il collegamento ad *authority file* ed enciclopedie online. Grazie all'adozione dei modelli RDF, FRBR e BIBFRAME (nella versione 1.0), milioni di dati bibliografici in formato MARC derivanti da tre ILS diversi (Aleph, Sebina e Millennium) e finora consultabili come registrazioni bibliografiche nei cataloghi online, in SHARE catalogue diventano condivisibili e modulari. La reale novità del modello di SHARE catalogue è la sperimentazione della tecnologia dei linked data usata per realizzare un catalogo cumulativo basato su dati bibliografici collegati, che in questo caso crea un'architettura di visualizzazione di secondo livello rispetto agli OPAC dei singoli atenei. Entrambe le strutture di visualizzazione dei cataloghi, quindi, convivono in SHARE catalogue: quella in dati collegati di tipo bibliografico costituisce un'architettura di visualizzazione e di ricerca cumulativa delle pubblicazioni possedute da tutti gli atenei coinvolti nel progetto; il catalogo cumulativo rinvia poi ai cataloghi online dei singoli atenei, ciascuno dei quali mantiene le proprie caratteristiche di ricerca e di visualizzazione delle registrazioni bibliografiche e del posseduto. Per quanto riguarda l'architettura del portale, unico punto di accesso al patrimonio delle risorse integrate, mediante un'interfaccia intuitiva è possibile ricercare informazioni a tre differenti livelli: persone-opere, istanze o pubblicazioni, item. Per agevolare la fruizione da parte degli utenti finali sono stati realizzati due *cluster* molto utili, ossia un indice cumulativo dei nomi e un punto di accesso autorizzato per i titoli, che consentono di effettuare una ricerca che è allo stesso tempo arricchita e semplificata (grazie al meccanismo che consente di riunire, sotto un unico identificativo, le forme varianti dei nomi degli autori, così come titoli diversi associati alle pubblicazioni presenti nei diversi cataloghi) rispetto a quella negli OPAC nativi.

Titia van der Werf (OCLC Research, Leiden) ha fornito un breve aggiornamento del lavoro che OCLC sta svolgendo nell'ambito del Web dei dati, in particolare per quanto riguarda l'estrazione delle entità in WorldCat a partire dai dati in MARC, le questioni legate alla pubblicazione, all'uso e alla visualizzazione delle risorse in linked open data, l'estensione del controllo di autorità e il potenziamento delle strutture bibliografiche multilingue³⁸. Si tratta di un percorso che sta muovendo i primi passi

36 SHARE (Scholarly Heritage and Access to Research) catalogue: <<http://catalogo.sharecat.unina.it/sharecat/clusters>>.

37 Cfr. Tiziana Possemato; Roberto Delle Donne, *SHARE catalogue: un'esperienza di cooperazione*, «Biblioteche oggi», 35 (2017), n. 1, p. 21-29, DOI: 10.3302/0392-8586-201701-021-1.

38 Cfr. Carol Jean Godby; Shenghui Wang; Jeffery K. Mixer, *Library linked data in the cloud: OCLC's experiments with new models of resource description*. San Rafael: Morgan & Claypool, 2015.

e che richiederà il contributo di diversi soggetti impegnati nell'implementazione dei linked open data nella comunità bibliotecaria in quanto «il Web dei dati non si organizza da sé».

L'evoluzione della descrizione archivistica ha seguito un percorso simile a quello della descrizione bibliografica, approdando anch'essa ai linked open data come strumento per l'esplorazione di una rete di relazioni sempre più ampia e l'incremento dei punti di accesso diretti alla fonte. In un intervento particolarmente efficace, Giovanni Michetti (Sapienza Università di Roma) ha discusso i benefici ma anche le criticità dei linked open data in ambito archivistico. I primi sono facilmente riconoscibili e incidono positivamente sull'usabilità, l'adattabilità e la ricercabilità dei dati. Tuttavia, l'adozione dei linked open data implica una notevole frammentazione dell'informazione e il conseguente rischio di perdere di vista il contesto generale, confondendolo nel fittissimo reticolo delle relazioni tra i dati. È parso dunque opportuno e motivato l'invito a bilanciare la granularità (che reca con sé innegabili vantaggi dal punto di vista della gestione del dato) con il dominio di riferimento, in particolare quando si tratta di lavorare con dati di archivio. Si assiste dunque a una sorta di paradosso, per cui da un lato il modello linked open data garantisce l'interoperabilità e orienta le istituzioni culturali verso una convergenza non solo tecnologica, dall'altro lato le singole comunità che compongono il variegato mondo GLAM hanno necessità di fare riferimento ai propri contesti specialistici, affini ma sostanzialmente diversi. L'analisi delle principali piattaforme (Linked open copac and archives hub³⁹, Europeana⁴⁰, LodLive⁴¹, ReLOAD⁴², CDEC⁴³, Archivio centrale dello Stato⁴⁴) ha evidenziato i problemi inevitabili nel passaggio dai modelli descrittivi tradizionali, caratterizzati dal testo libero e dal carattere narrativo, ai linked open data, in termini di stabilità, autorialità e neutralità. Non bisognerebbe, insomma, dimenticare le parole di Jacques Derrida, secondo il quale la struttura tecnica dell'archivio determina la struttura del suo contenuto, e impegnarsi nella ricerca di un equilibrio non facile tra la granularità del dato (che reca con sé innegabili vantaggi) e il dominio di riferimento, in particolare quando si lavora su fonti archivistiche⁴⁵.

Enzo Borsellino (Università Roma Tre) ha ricostruito le ultime tappe della riorganizzazione del MiBACT attraverso un'analisi dei recenti decreti legislativi, soffermandosi in particolare sulla questione dell'autonomia dei musei, o meglio di un ristretto numero di musei. Borsellino ha individuato nell'assenza di un disegno organico tra musei, archivi e biblioteche il principale problema della cosiddetta riforma Franceschini, che sembra porsi l'unico obiettivo di incrementare il numero dei visitatori nei musei più grandi e famosi, trascurando la specificità del concetto di 'museo diffuso' delle collezioni museali italiane. La diminuzione degli organici e dei ruoli dirigenziali e l'assenza di investimenti nel potenziamento dell'azione di tutela sul

39 LOCAH – Linked archives hub dataset: <<http://data.archiveshub.ac.uk/>>.

40 Europeana: <<http://www.europeana.eu/>>.

41 LodLive: <<http://lodlive.it/>>.

42 ReLoad: <<http://labs.regesta.com/progettoReload/>>

43 CDEC dati: <<http://dati.cdec.it/>>.

44 Dati.ACS: <<http://dati.acs.beniculturali.it/>>.

45 Cfr. Jacques Derrida, *Mal d'archive: une impression freudienne*. Paris: Galilée, 1995 (traduzione italiana: *Mal d'archivio: un'impressione freudiana*. Napoli: Filema, 1996).

territorio e nei musei con personale adeguato alle necessità reali di un patrimonio così vasto, hanno spinto circa un migliaio di operatori del settore (tra cui funzionari, restauratori, archeologi, storici dell'arte, demotnoantropologi, museologi, direttori di museo, ex soprintendenti) a manifestare il 7 maggio 2016 contro la separazione fra tutela e valorizzazione, l'accorpamento delle soprintendenze e l'inadeguatezza delle risorse economiche, ovvero le principali criticità della riforma.

Luca Martinelli (Wikimedia Italia), ha presentato le principali caratteristiche di una piattaforma più volte evocata durante il convegno, ossia Wikidata, una base di conoscenza liberamente utilizzabile e modificabile da umani e computer, pubblicata con una licenza libera, che permette il riutilizzo dei dati per qualunque scopo, anche commerciale⁴⁶. Gli scopi principali di questo database secondario, libero, collaborativo e multilingue sono i seguenti: centralizzare l'accesso e la gestione dei dati strutturati delle voci di Wikipedia e dei suoi progetti fratelli, raccogliere i dati fondamentali delle voci, creare delle *query* automatiche, rendere i dati riutilizzabili (anche da terzi). I dati sono strutturati in 'dichiarazioni', ossia coppie 'proprietà-valore' (che possono contenere qualificatori opzionali) e la fonte da cui è stato ricavato il dato. Lanciata ufficialmente il 30 ottobre 2012, Wikidata comprende oltre 25 milioni di elementi, è gestita da una comunità di volontari, che decide le linee guida per la creazione e la manutenzione dei dati, e ha avviato forme di collaborazione con importanti istituzioni del mondo GLAM, come la Biblioteca nazionale centrale di Firenze, l'Istituto centrale per il catalogo unico, il MoMA di New York, la Tate Gallery, il Rijksmuseum o lo Human genome project⁴⁷. Tra i nuovi progetti sviluppati sulla base di Wikidata, merita di essere menzionato WikiCite, che ha l'obiettivo di inserire, mediante la proprietà 'cita', libri e articoli scientifici su Wikidata e creare così un network citazionale fra pubblicazioni ovvero un database centralizzato di dati bibliografici in forma strutturata, coniugando l'esperienza degli istituti culturali nella modellazione dei dati bibliografici (come FRBR) con l'approccio pragmatico del wiki⁴⁸.

Conclusioni

Il Web semantico e i linked open data hanno indotto, dunque, una radicale trasformazione degli strumenti di mediazione bibliografica, della relazione tra l'utenza e le risorse documentarie, del rapporto tra i cataloghi delle biblioteche e il più vasto ecosistema informativo digitale. Non sorprende che questo rapido mutamento abbia provocato un certo disorientamento in una comunità professionale ancora in gran parte legata al precedente paradigma teorico-pratico; eppure non è difficile scorgere in questa evoluzione tecnologica una grande opportunità per riaffermare in un contesto molto più ampio della propria comunità di riferimento il ruolo delle biblioteche e dei bibliotecari come produttori di informazioni e dati bibliografici di qualità, pronti per essere condivisi e integrati in rete.

Le biblioteche si muovono già da tempo sul terreno della cooperazione e della condivisione (di risorse e servizi), tuttavia i paradigmi dell'apertura, della convergenza e dell'integrazione di risorse e dati culturali eterogenei prefigurano uno scenario inedito, che va ulteriormente definito dal punto di vista teorico per poter esse-

⁴⁶ Wikidata: <<https://www.wikidata.org/wiki/>>.

⁴⁷ Cfr. Luca Martinelli, *Wikidata: la soluzione wikimediana ai linked open data*, «AIB studi», 56 (2016), n. 1, p. 75-85. DOI: 10.2426/aibstudi-11434.

⁴⁸ WikiCite: <<https://meta.wikimedia.org/wiki/WikiCite>>.

re applicato nei contesti più vari, tenendo sempre presente l'esigenza dell'interoperabilità. Il "Convegno AIB CILW 2016" segna senza dubbio una tappa importante di questo processo di maturazione concettuale e di restituzione di buone prassi, in un ambito non solo nazionale. Due prospettive appaiono particolarmente promettenti: il contributo che la comunità interdisciplinare dei professionisti della cultura può offrire per l'elaborazione di rappresentazioni sempre più efficaci dei dati e dei collegamenti tra loro, e la collaborazione, intensificatasi negli ultimi anni, tra il mondo GLAM e le piattaforme Wikimedia, il cui modello aperto e collaborativo è diventato – come hanno certificato quasi tutte le relazioni presentate durante il convegno – un punto di riferimento per i fornitori di sistemi documentari e per quanti si impegnano per la diffusione della conoscenza e il progresso culturale della società.

Articolo proposto il 27 febbraio 2017 e accettato l'8 marzo 2017.

ABSTRACT

AIB studi, vol. 57 n. 1 (gennaio/aprile 2017), p. 167-178. DOI 10.2426/aibstudi-11614

LUIGI CATALANI, Biblioteca provinciale, Potenza; Università degli studi di Salerno, Dipartimento di scienze del patrimonio culturale, Fisciano, e-mail lcatalani@unisa.it.

La convergenza delle risorse informative nel Web dei dati e delle relazioni (note sul Convegno AIB CILW 2016, Biblioteca nazionale centrale di Roma, 21 ottobre 2016)

Il contributo riporta i contenuti discussi nell'ambito del "Convegno AIB CILW 2016" intitolato "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, svoltosi il 21 ottobre 2016 presso la Biblioteca nazionale centrale di Roma. La riflessione corale sugli aspetti teorici e pratici della trasformazione delle modalità di gestione delle risorse informative prodotta dall'adozione del Web semantico e dei linked open data, ha consentito di apprezzare alcune tra le buone pratiche messe in opera dai professionisti del settore culturale e di mettere a fuoco alcune prospettive particolarmente promettenti, quali lo sviluppo di nuovi modelli di visualizzazione dei dati e l'integrazione dei dati prodotti dagli specialisti GLAM con quelli prodotti dalla comunità wikimediana.

The convergence of information resources on the Web of data and relations (remarks on the AIB CILW 2016 Conference, National Central Library of Rome, October 21, 2016)

The paper reports the "AIB CILW 2016 Conference" entitled "Revamping information resources: granularity, interoperability and data integration", organized by the AIB study group on Cataloguing, indexing, linked open data and semantic web, held on October 21, 2016, at the National Central Library, Rome. Choral reflection on the theoretical and practical aspects of the transformation of the management of information resources produced by the adoption of the semantic Web and linked open data, has enabled the audience to appreciate some good practices put in place by the cultural professionals and to focus on some particularly promising prospects, such as the development of new data visualization models and the integration of data produced by GLAM specialists with those produced by the Wikimedia community.