

Linked data nelle biblioteche digitali e di ricerca

VALERIA LO CASTRO

Università degli studi di Napoli "Federico II"
valeria.locastro@unina.it

Una panoramica internazionale

*Se vuoi andare veloce, vai da solo.
Se vuoi andare lontano, vai insieme.¹*

Introduzione

Istituzioni culturali come le università e le biblioteche hanno da sempre raccolto dati relativi alla loro attività gestionale, amministrativa, didattica e di ricerca (per esempio tutte le informazioni relative ai dipartimenti, ai docenti, ai moduli dei corsi, ai progetti di ricerca ecc.), e dati relativi al loro posseduto (libri, riviste, ebook, *e-journals*, banche dati, risorse didattiche ecc.). Si tratta di informazione spesso altamente strutturata, contenuta in diversi ambienti o silos.

Questa funzione si è accentuata con l'aumento delle informazioni e la nascita delle biblioteche digitali. Possono essere ascritti al campo delle biblioteche digitali – per le quali non esiste ancora una definizione univoca e condivisa – gli archivi della ricerca, le piattaforme di *e-learning*, gli archivi delle tesi di dottorato e delle riviste open access ecc.²

In molti casi le biblioteche e i bibliotecari hanno contribuito alla creazione di questi depositi istituzionali, mettendo a frutto le competenze acquisite in relazione allo sviluppo delle collezioni, alla gestione dei metadati e dei servizi di indicizzazione e recupero dell'informazione.

In anni recenti gli standard e gli strumenti del Semantic Web e le tecnologie linked open data hanno rappresentato il tentativo di creare un linguaggio uniforme e condiviso per la descrizione di dati e metadati di diverso tipo. Attraverso l'utilizzo di standard e di un linguaggio comune a tutte le comunità presenti sul web, è possibile raggiungere l'interoperabilità tra i vari sistemi, che possono così superare l'isolamento. Tale linguaggio è inoltre comprensibile non soltanto agli umani, ma anche – e soprattutto – alle macchine, che possono così produrre anche dei meccanismi di inferenza.³

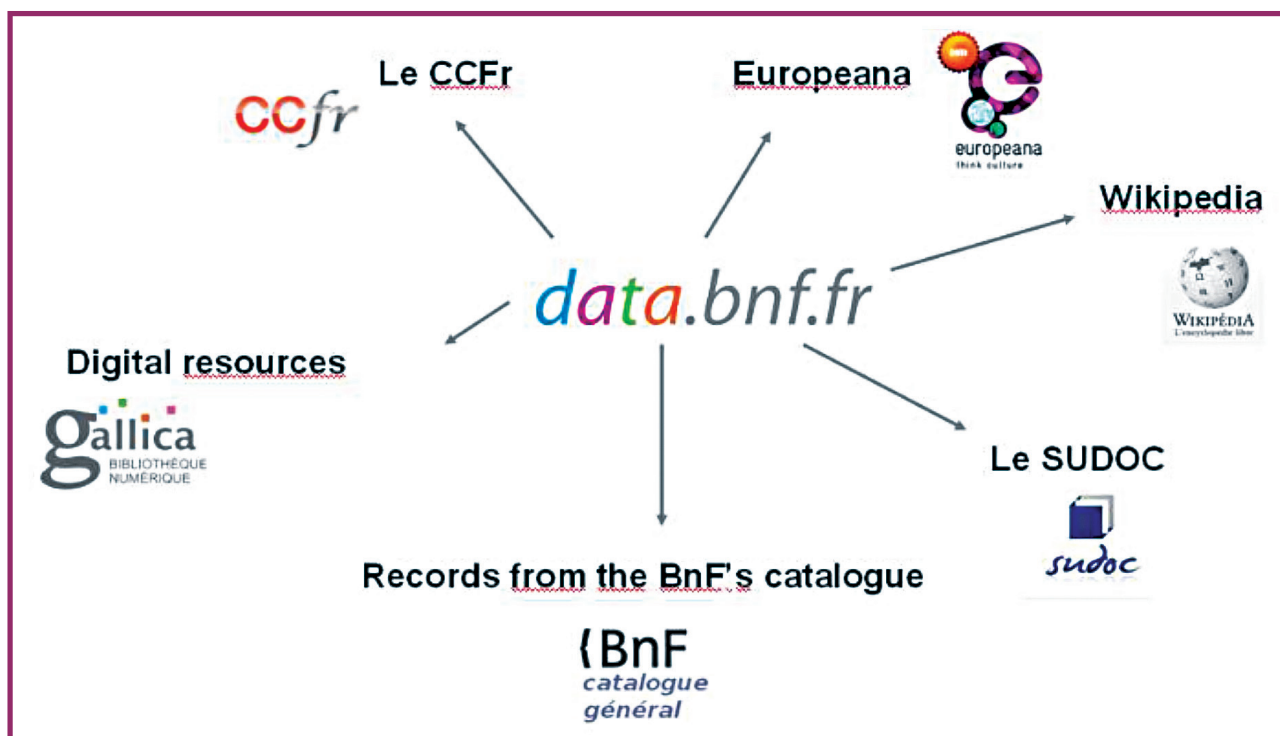
Il Semantic Web, nato dalla necessità di mettere ordine in un ecosistema informativo caratterizzato da grandi quantità di dati e per aumentare le potenzialità del web tradizionale (e superare alcune criticità relative alla ricerca di documenti e informazioni), e la sua versione "debole" costituita dai linked (open) data (LOD)⁴ vanno a costituire un universo che si compone di alcuni elementi essenziali e imprescindibili:

- Uniform Resource Identifier (URI) per identificare univocamente una risorsa;
- Resource Description Framework (RDF) come modello logico per descrivere i collegamenti e le relazioni tra i metadati e le risorse;
- vocabolari e ontologie che definiscono i concetti e i termini usati per descrivere e rappresentare una particolare area di interesse. Il loro scopo è quello di classificare i termini che vengono usati in un particolare settore disciplinare e indicare le relazioni e gli utilizzi del termine in un dato contesto (le ontologie presentano, a differenza dei vocabolari, un formalismo più stretto).

Negli ultimi anni le biblioteche hanno progressivamente riconosciuto questi modelli come adeguati ai propri contenuti, e ne hanno compreso le potenzialità nel favorire una diffusione molto più ampia di dati e metadati bibliografici standardizzati, nonché il loro riutilizzo in contesti anche diversi da quelli di origine e nel garantire una loro maggiore visibilità sul web.

Il collegamento tra materiale eterogeneo, proveniente da diversi contesti disciplinari consente di rompere l'isolamento dei silos entro i quali è contenuta l'informazione e di costruire percorsi di ricerca personalizzati e vere e proprie voci di enciclopedia.

In particolare, attraverso l'uso di principi, tecnologie e *datasets* disponibili in linked open data, i contenuti di un'università possono essere interconnessi con i relativi materiali di studio e di ricerca nella singola università e



nel resto del mondo. Per esempio, una pagina web che descrive un corso può essere interconnessa automaticamente con i correlati risultati di ricerca, pubblicazioni, progetti, pagine di Wikipedia, gruppi di ricerca, ricercatori, conferenze video, materiali didattici aperti e così via. In questo contributo vengono presentati tre progetti nel campo dei Library Linked Data.⁵ Il primo è uno dei più maturi in questo settore, vale a dire quello di “data.bnf.fr” della Bibliothèque nationale de France (BnF). Proprio perché si tratta di un progetto più consolidato, si è scelto in questo caso di soffermarsi più che sulle scelte tecniche, sui risultati raggiunti in termini di servizi e uso rispetto agli obiettivi iniziali. Verrà poi presentato un recentissimo progetto (Linked Data for Libraries) che vede coinvolte tre grandi biblioteche di università americane (Harvard, Stanford, Cornell) nel collegamento attraverso i linked data di risorse accademiche di diverso tipo. Infine verrà raccontata un’analoga esperienza europea, guidata però in questo caso da tecnici e informatici di diverse università. Questi ultimi due casi sono ancora in una prima fase di sviluppo, per questo motivo verranno indagate primariamente le finalità, nonché le scelte tecniche e tecnologiche. L’aspetto innovativo, in questi ultimi due casi – e che certamente riguarda le biblioteche accademiche – è dato dalla maturata consapevolezza da un lato di dover creare un ambiente che colleghi le informazioni relati-

ve al posseduto e quelle relative alla comunità di riferimento, dall’altro di attivare sempre più meccanismi di cooperazione per esplicitare totalmente il potenziale dei linked open data e per rendere i dati sempre più fruibili. La cooperazione diventa dunque l’altra parola chiave, insieme a riuso e interoperabilità.

data.bnf.fr

data.bnf.fr⁶ è la piattaforma linked open data sviluppata e gestita dalla Bibliothèque nationale de France. Essa ha come scopo finale quello di coprire tutti i cataloghi e le risorse della BnF, ma è stata inizialmente implementata in funzione del progetto Gallica per renderne le risorse più visibili.

Il modello di dati e il lavoro di specifica è iniziato nel 2009, lo sviluppo effettivo nel novembre del 2010, e la prima versione è andata online nel luglio del 2011. L’interfaccia web riunisce i dati, i collegamenti e i legami provenienti da vari silos sia interni (il suo catalogo generale e il catalogo di archivi e manoscritti, Gallica ecc.) sia esterni (il catalogo collettivo delle biblioteche di ricerca francesi SUDOC, il catalogo unico francese CCFR, WorldCat, VIAF, Wikipedia ecc.).

L’obiettivo finale è quello di consentire agli utenti di trovare sia le informazioni relative a una registrazione bibliografica, sia il documento completo se esiste, sia

Bienvenue sur data.bnf.fr

La Bibliothèque nationale de France vous guide dans ses ressources en regroupant sur une même page toutes les informations issues de ses différents catalogues, ainsi que de sa bibliothèque numérique Gallica.

Le projet utilise les outils du Web sémantique et s'inscrit dans une démarche d'ouverture des données.

Enquête de l'Observatoire des publics de la BnF : [Votre avis nous intéresse](#)

Auteurs les plus consultés

- Voltaire (1694-1778)
- Tacite (00557-01207)
- Alexandre Dumas (1802-1870)
- Érasme (1469-1536)
- Claude Galien (01317-02017)

Œuvres les plus consultées

- Les trois mousquetaires
- Les travailleurs de la mer
- L'amitié
- Chemin de longue étude
- Fables

Regard sur

Platon (04277-0348? av. J.-C.)

il collegamento ad altre risorse esterne. La BnF, pur essendo impegnata in una riflessione internazionale su nuovi modelli bibliografici, schemi e regole di catalogazione e standard come Functional Requirements for Bibliographic Records (FRBR) e Resource Description and Access (RDA), nello stesso tempo attraverso questo progetto inizia a confrontarsi con le possibilità di utilizzo dei dati e con le implicazioni pratiche, le opportunità e le criticità correlate.

Dal punto di vista tecnico i dati dei cataloghi (in formato Machine Readable Cataloguing, MARC), degli archivi (Encoded Archival Description, EAD) e delle risorse digitali (Dublin Core, DC) sono estratti e trasformati in RDF. La Bibliothèque nationale de France usa identificativi persistenti per web URIs: gli identificativi Archival Resource Key (ARK), che vengono usati per identificare record del catalogo, risorse archivistiche, oggetti digitali da Gallica e authority record. Il modello di dati è stato costruito a partire dal modello FRBR (CubicWeb è il software gratuito utilizzato per creare le corrispondenze e i collegamenti automatici). Gli autori e i soggetti sono collegati ad authority file esterni come AGROVOC,⁷ Geonames, DBpedia, Wikipedia, Virtual International Authority File (VIAF), Dewey, Library of Congress e Deutsche Nationalbibliothek. I vocabolari utilizzati sono Simple

Knowledge Organization System (SKOS),⁸ Friend of a Friend (FOAF),⁹ DC e RDA.¹⁰

La piattaforma genera pagine web che forniscono informazioni standardizzate, riferimenti e collegamenti con gli autori, opere e soggetti che gli utenti possono trovare attraverso i motori di ricerca e da cui possono navigare.

Queste pagine riuniscono dati descrittivi organizzati in un modo conforme ai requisiti e alle entità FRBR. Per un determinato autore si trovano tutti i link alle pagine web delle opere rilevanti (di e su un autore, in due differenti sezioni); per un'opera, un link alla pagina del suo autore, ma anche a tutte le altre manifestazioni del lavoro (risorse bibliografiche, materiale online). Tutti i datasets grezzi sono disponibili per il libero download in formato RDF.

Le pagine costruite da data.bnf.fr non sono destinate a sostituire i cataloghi esistenti e gli altri silos che essi sfruttano, ma a fornire un collegamento. In breve, sono finalizzate a rendere i dati della biblioteca meglio funzionanti sul web, una sorta di hub di informazioni bibliografiche, con dati espliciti strutturati e URIs permanenti, che funzioni come un ambiente sicuro di dati affidabili. Queste pagine web presentano una serie di funzionalità avanzate (filtra per tipo, esporta e invia, citazione sui social network). I dati possono essere anche recupera-

ti automaticamente da altri datasets open in modo da migliorare la corrispondenza, e fornire informazioni da questi insiemi.

Nel 2011, al fine di soddisfare i requisiti completi di LOD e di essere open non solo tecnicamente ma anche giuridicamente, la BnF ha adottato licenze pubbliche aperte raccomandate dal French Government Open Data (Etalab) per il suo programma data.gouv.fr per i propri metadati pubblicati in RDF. Queste licenze possono essere paragonate a un tipo di licenza CC-BY (gratuita, aperta, favorisce il riuso e la riproduzione anche a fini commerciali), compatibile con altre licenze straniere internazionali simili, come Open Government License (UK Government) e le licenze ODC-BY e CC-BY 2.0.¹¹ Alla fine del 2012, 200.000 pagine web sono state rese disponibili su data.bnf.fr. Queste pagine descrivono 45.250 opere e 20.000 autori (entità FRBR), legati a 3,6 milioni di record bibliografici. Si stima che le informazioni presenti su data.bnf.fr rappresentino il 20% del contenuto disponibile nel catalogo BnF. Durante il suo primo anno completo di funzionamento, data.bnf.fr ha accumulato 637.650 visitatori individuali e 1,2 milioni di visite alle pagine (+550% per visitatori e +400% per visita a pagine). Su base mensile, questo significa una media di 53.134 visitatori per mese nel 2012.

È significativo rilevare che l'80,6% delle visite provengono da un motore di ricerca web. Questo conferma che la maggior parte degli utenti del web utilizza motori di ricerca generali per qualunque tipo di bisogno, compresi quelli di natura bibliografica. Appare dunque evidente l'importanza che queste piattaforme (come data.bnf.fr), a differenza di molti OPAC e database tradizionali, siano raggiungibili in maniera trasparente attraverso i motori di ricerca. Inoltre tale condizione fa sì che utenti originariamente non interessati ai dati bibliografici possano scoprire contenuti di loro interesse partendo da ricerche di carattere generale.

Inoltre il 71,9% delle visite per data.bnf.fr ha portato alla visita di altri ambienti della BnF (come i cataloghi o Gallica). Questo sembrerebbe confermare che data.bnf.fr gioca completamente il suo ruolo di informazione "hub", stimolando la visita ad altre risorse BnF.

A luglio 2014 si è raggiunto il numero di 400.000 autori collegati a oltre 7 milioni di documenti provenienti da BnF catalogue général e da BnF archives et manuscrits. È difficile individuare e seguire tutti i tipi di riuso realizzati. La BnF prevede che nel giro di 5 o 10 anni le biblioteche pubbliche in Francia (e altrove) potrebbero utilizzare le tecnologie del Semantic web per col-

legare i loro cataloghi locali ai metadati di riferimento prodotti dalla Bibliographic National Agency of France della stessa biblioteca invece che continuare il loro recupero attraverso il protocollo Z39.50, i server FTP o i repositories OAI. Presupponendo che la maggior parte dei riferimenti che descrivono le collezioni conservate da queste biblioteche dovrebbero essere rese disponibili dalla BnF in un hub di metadati costruito sulle tecnologie sviluppate per data.bnf.fr (che sono tutte tecnologie open source), le biblioteche locali potrebbero usare appropriate Application Programming Interface (API)¹² per collegare i metadati della BnF dai loro sistemi locali invece che recuperarli. Questo potrebbe portare una seconda vita digitale alla French National Bibliography e aumentare considerevolmente l'utilizzo e il riutilizzo di metadati.

In questo scenario, la catalogazione locale potrebbe essere ridotta al minimo, consentendo ai bibliotecari di reinventare i loro compiti descrittivi, concentrandosi più sul collegamento tra risorse e l'arricchimento di metadati con le informazioni locali, quando sono utili, invece che sulla descrizione.

Sul versante internazionale il progetto francese coopera con altre biblioteche nazionali impegnate in LOD, come le biblioteche di Germania, Svezia, la British Library o la Library of Congress o tramite altri progetti multilaterali come VIAF.

Le risorse della biblioteca sono utilizzate anche da altre comunità oltre quella dei bibliotecari: l'agenzia culturale del French Foreign Affairs ha già usato i dati di data.bnf.fr per sviluppare IF Verso, un portale che funge da registro per tutte le opere francesi tradotte in lingue straniere, una start-up ha creato un'applicazione per iPhone che usa i dati e collega i metadati su autori (prevenienti da data.bnf.fr), opere digitalizzate (alcune provenienti da Gallica) e informazioni geospaziali (allineate con i nomi di autorità geografici della BnF) in modo da sviluppare un nuovo servizio dove i turisti e gli abitanti di tutte le regioni della Francia possano facilmente leggere sul loro smartphone estratti da libri digitalizzati in relazione al monumento o al luogo in cui si trovano.

Risulta evidente che l'obiettivo di lungo periodo riguarda la scalabilità dei dati (si prevede entro tre anni di esporre tutti i metadati di BnF in data.bnf.fr), il monitoraggio del volume dei dati esposti, la scelta dei dati da esporre. Si fa inoltre via via più pressante l'esigenza di un coinvolgimento sempre più forte di persone provenienti da diversi ambiti professionali.

Linked Data for Libraries Project

Cornell University Library, Harvard Library Innovation Lab e Stanford University Libraries hanno ricevuto, il 5 dicembre 2013, dalla Fondazione Andrew W. Mellon un finanziamento per due anni di un milione di dollari per la realizzazione del Linked Data for Libraries Project (LD4L),¹³ il cui obiettivo è quello di creare un sistema che estrae le informazioni dai diversi silos esistenti in un formato comune sviluppando un'ontologia e dei linked data, ossia metadati, relazioni tra dati e un ampio contesto per migliorare la scoperta e l'accesso alle informazioni accademiche delle tre biblioteche.

“Attualmente, un sacco di informazioni su libri, articoli e materiale culturale che potrebbe essere più facile da trovare e capire è nascosto in sistemi sparsi di molte biblioteche”, dice Dean Krafft, Library Chief Technology Strategist di Cornell. “Noi abbiamo l'opportunità di usare i linked data come un formato comune per riunire insieme l'informazione sparsa. Alla Cornell, abbiamo lavorato con il Semantic Web per più di un decennio, e siamo molto lieti che la Mellon Foundation ci dia l'opportunità di approfondire il lavoro con i nostri colleghi di Harvard e Stanford”.¹⁴

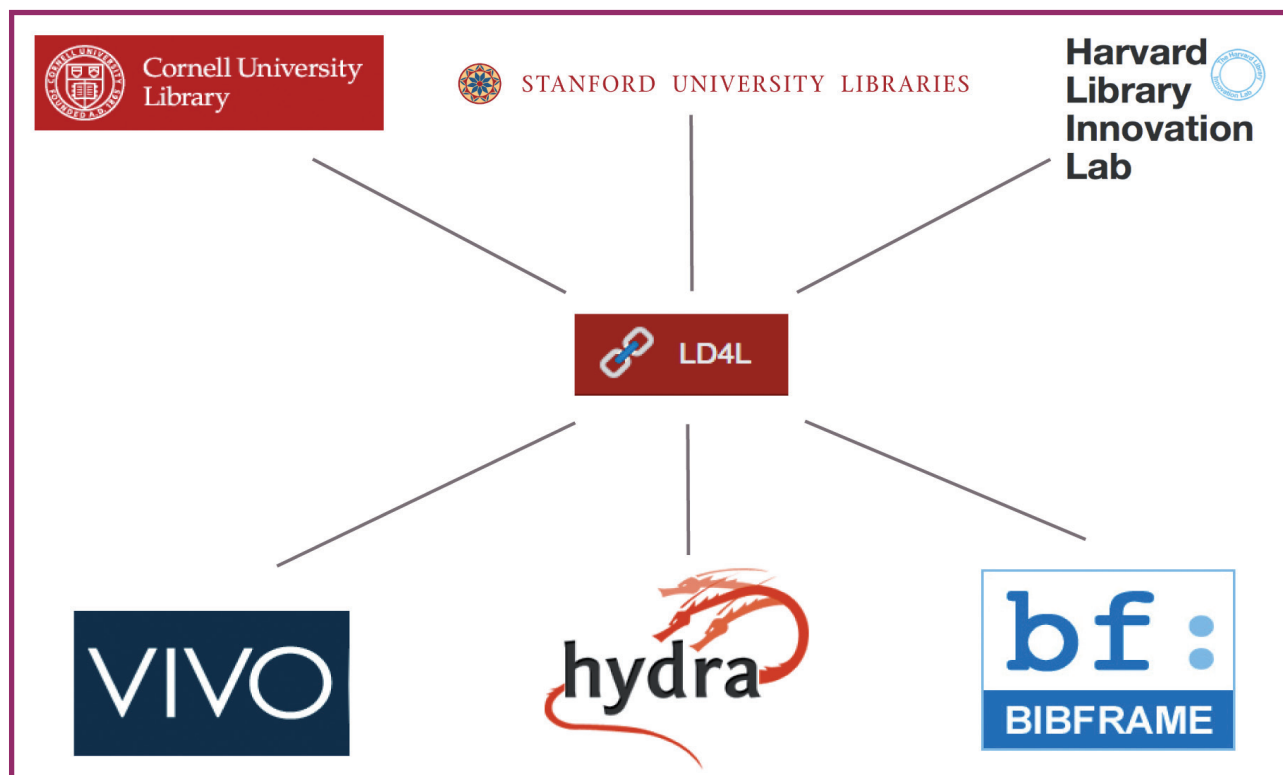
“Il nostro laboratorio è entusiasta all'idea di lavorare con partner leader su un progetto che renderà più faci-

le per le biblioteche condividere in maniera aperta sempre di più di ciò che sanno – non solo sul loro contenuto, ma su come viene utilizzato quel contenuto”,¹⁵ dice David Weinberger, co-direttore di Harvard Library Innovation Lab.

“Una delle prospettive interessanti per questo progetto è che si *testerà* l'usabilità e l'utilità dei dati prodotti. Molti elementi di progetto si sovrappongono al Technology Plan sviluppato come parte dello Stanford Linked Data Workshop nel 2011 e ci offrono l'opportunità per una grande collaborazione tra i tre sistemi bibliotecari”, conclude Tom Cramer, Chief Technology Strategist per le Stanford University Libraries.¹⁶

Durante un meeting svoltosi alla Cornell University l'8 e 9 dicembre 2014, sono stati presentati i risultati del primo anno di lavoro e gli obiettivi per il 2015. Il primo anno di lavoro è stato dedicato alla progettazione dell'ontologia, alla scelta del modello di dati, all'identificazione dei *datasets* da esporre e dei vocabolari esterni.

L'obiettivo è creare un modello di archivio di risorse informative semantiche accademiche (Scholarly Resource Semantic Information Store model) che funzioni sia all'interno di istituzioni individuali sia attraverso una rete coordinata ed estensibile di LOD, allo scopo di catturare il valore intellettuale che i bibliotecari e gli altri esperti di dominio aggiungono alle risorse informative



quando essi descrivono, annotano, organizzano, selezionano e usano queste risorse, unitamente al valore sociale testimoniato dai modelli di utilizzo.¹⁷

I tre partner hanno deciso inoltre di utilizzare le tecnologie già esistenti e di sfruttare sia il lavoro fatto dalle stesse istituzioni negli anni precedenti che quello di altri, aderendo così alla filosofia del riuso propria dei progetti di questo tipo.

Per quanto riguarda la costruzione dell'ontologia si è deciso di adottare VIVO,¹⁸ uno strumento per la rappresentazione delle informazioni sulla ricerca e i ricercatori (i lavori accademici, gli interessi di ricerca, le relazioni organizzative) già sviluppato nel 2003 dalla stessa Cornell University. VIVO è un'ontologia standard che interconnette i ricercatori e le comunità utilizzando linked open data, ma anche un dataset e una comunità aperta con una forte partecipazione nazionale e internazionale (alla fine del 2012 oltre venti paesi e cinquanta organizzazioni hanno utilizzato VIVO per connettere ricercatori, pubblicazioni, risorse per la ricerca, finanziamenti, corsi tenuti e altre attività scientifiche).

Il modello di dati è invece BIBFRAME,¹⁹ un progetto della Library of Congress per la creazione di un modello di dati standard con il quale le biblioteche possono descrivere e scambiare informazioni su risorse accademiche.

Si è scelto di partire dai dati esistenti dei cataloghi delle biblioteche in MARC e convertirli in RDF e sfruttare il lavoro di OCLC²⁰ sugli identificatori bibliografici per agganciarli a quelli delle tre biblioteche. Harvard ha l'82% di 13,6 milioni di record bibliografici che possono essere agganciati al lavoro sugli identificativi di OCLC; Stanford e Cornell hanno 2,6 milioni di record comuni su 5,8 milioni. Fin qui è stato completato il lavoro inerente ai dati provenienti da Cornell. I vocabolari scelti, al momento, sono SKOS e Getty.

Altro elemento su cui fa perno l'intero progetto è HYDRA,²¹ un software ricco di funzioni e una comunità che si occupa di creare archivi digitali e collezioni, fondata nel 2008 da Stanford University, University of Virginia, University of Hull e Fedora Commons.

Gli obiettivi del 2015 riguardano l'esposizione dei dati di Harvard e Stanford, lo sviluppo della versione beta di un motore di ricerca basato su RDF, il rilascio pubblico del codice sorgente del progetto e dell'ontologia.

Verrà inoltre organizzato un workshop, da tenersi a febbraio, al quale saranno invitate venticinque persone, provenienti da dodici diverse istituzioni tra biblioteche,

archivi e istituzioni della memoria, al fine di mostrare il prototipo LD4L per ottenere un feedback sulla progettazione dell'ontologia, sul design e l'approccio e per spronare altre istituzioni a seguire questa strada. È interessante notare che queste tre grandi istituzioni hanno riaccolto attorno a questo progetto diverse figure professionali (ricercatori, sviluppatori, produttori di metadati, catalogatori), cercando di aderire alla politica del riuso tipica dei linked data.

Anche qui, come alla BnF, si è cominciato a lavorare pur sapendo di non poter raggiungere un risultato perfetto e consapevoli di doversi confrontare con una serie di criticità.

Linked Universities

Questo progetto è guidato da un gruppo di ricercatori universitari provenienti dall'area informatica che si avvale anche di competenze bibliotecarie.

Linked Universities²² è un'alleanza di università europee impegnate nell'esposizione dei loro dati pubblici in linked data. Sono coinvolte università dell'area anglosassone (University of Bristol, University of Southampton, The Open University), tedesche (University of Münster), finlandesi (Aalto University), spagnole (Universitat Pompeu Fabra), turche (Ege University), greche (Aristotle University of Thessaloniki) e slovacche (Charles University). Queste iniziative in campo accademico sono ancora sporadiche e spesso scollegate tra loro (con un'enorme dispersione di energia in fase di sviluppo di singole piattaforme e di processi che vengono spesso duplicati): Linked Universities nasce per raccordare e collegare i diversi progetti.

L'idea di fondo è che il potenziale dei linked data nel campo della didattica e della ricerca vada ben oltre il beneficio individuale per la singola istituzione, e che tale potenziale possa essere pienamente esplicito solo attraverso la realizzazione di reti di dati universitari (*web of university data*), realizzate aggregando, integrando e comparando dati provenienti da diverse università, la condivisione di pratiche e la collaborazione nello sviluppo di piattaforme.

Il portale linkeduniversities.org è uno spazio collaborativo, nel quale le istituzioni e gli individui coinvolti nell'esposizione di linked data universitari possono descrivere, condividere e riutilizzare *datasets*, vocabolari, tool e pratiche comuni.

In esso è possibile scaricare sia i *datasets* delle singole università coinvolte (interrogabili attraverso l'*endpoint*

SPARQL) che contengono dati (in RDF) relativi a ciò che l'università pubblica, come per esempio cataloghi di biblioteche, informazioni sui corsi online, podcast, informazioni sugli archivi della ricerca, video, Open Educational Resources (OER) ecc., sia informazioni relative alla sua organizzazione e alla comunità di riferimento, come informazioni sul personale, sugli edifici, eventi e news.

Sono presenti modelli standardizzati di metadati per supportare la descrizione e lo scambio relativo a informazioni sui corsi; vocabolari per la descrizione dell'organizzazione interna dell'università, vocabolari per le pubblicazioni accademiche e le comunità scientifiche (come BIBO - The Bibliographic Ontology e la già citata VIVO); vocabolari per la descrizione di risorse multimediali e per i progetti di ricerca.

Ci sono poi diversi tipi di strumenti (*tools*) per la conversione in RDF dei dati, dei cataloghi delle biblioteche, dei *repositories* open archive, dei riferimenti bibliografici, dei record MARC, delle risorse didattiche. Infine esistono strumenti per la gestione e condivisione di comunità come VIVO ed ePrints (una piattaforma per gestire i *repositories* della ricerca nelle università).

Tra le varie università coinvolte particolarmente significativo è l'apporto di The Open University, i cui ricer-

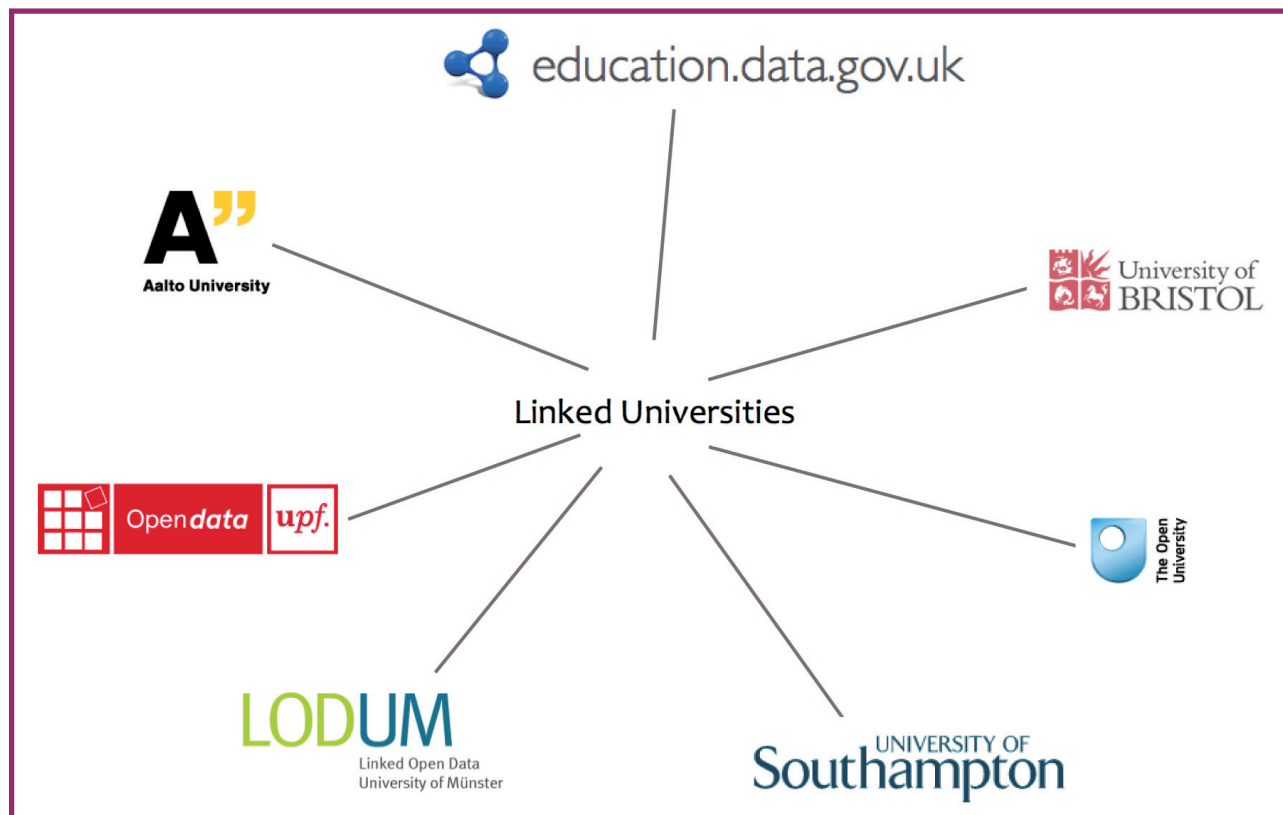
catori sono l'anima e la mente del progetto, nonché i principali fornitori di vocabolari e *tools*.

The Open University²³ è la più grande università della Gran Bretagna: creata nel 1969, ha 250.000 studenti per anno, 8.000 docenti, un grande campus a Milton Keynes, quasi tutti i corsi fruiti open e a distanza, 13 centri regionali, più centri nazionali, corsi disponibili in un gran numero di paesi.

Il Knowledge Media Institute, il principale centro di ricerca sulle tecnologie Semantic web dell'Ateneo, al cui interno lavorano 80 tra ricercatori, dottorandi, associati, ha sviluppato data.open.ac.uk., la prima piattaforma linked open data che contiene informazioni aperte provenienti da un intero ateneo.

Sono moltissime le tipologie di dati presenti: 580 moduli di descrizioni dei corsi, 16.000 articoli accademici, 220 video podcast e 1.500 audio podcast, 640 Open Learn Units OER, 900 video sul canale Youtube, 100 tra mappe e immagini di edifici dell'università, 12.000 libri del catalogo della biblioteca.

I dati relativi a queste risorse sono stati convertiti in formato RDF, mentre è stato scelto il modello MLO - Metadata for Learning Opportunities per supportare la descrizione, lo scambio relativo a informazioni sui corsi e VoID data model²⁴ supportati da vocabolari e



ontologie come FOAF per le persone e le istituzioni, Dublin Core e BIBO per le risorse accademiche e altri vocabolari come RDF, OWL e SKOS (in particolare per rappresentare i luoghi).

Questa piattaforma è collegata, oltre che alle istituzioni della rete Linked Universities, anche a mEducator²⁵ e OrganicEduNet²⁶ (due progetti europei che aggregano risorse didattiche) e Data.gov.uk Education,²⁷ che aggrega informazioni educative in Gran Bretagna.

Sulla piattaforma data.open.ac.uk è possibile trovare i collegamenti alle diverse tipologie di informazioni, interrogare i dati attraverso l'endpoint SPARQL, conoscere le nuove app nate dall'utilizzo di questi dati, studiare con il servizio di connessione in mobilità, supportare la valutazione della ricerca ecc.²⁸

Conclusioni

La breve rassegna presentata in questo contributo solleva alcune questioni su cui biblioteche e bibliotecari sono chiamati a riflettere.

Intanto la necessità di costruire, nel momento in cui si sceglie di dar vita a progetti di questo tipo, gruppi di lavoro con diverse competenze. In alcuni casi, come abbiamo visto, sono gli informatici a guidare i processi, in altri i bibliotecari. In ogni caso l'apporto di questi ultimi risulta prezioso nella misura in cui essi possono mettere a disposizione le loro competenze per garantire la qualità e l'autorevolezza dei *datasets* e dei vocabolari.

Anche il lavoro di descrizione che fin qui ha notevolmente impegnato i bibliotecari potrebbe diminuire in favore dei processi di identificazione e collegamento.²⁹

Portare i propri dati strutturati in questo mondo significa valorizzarli e dunque in qualche modo mettere a frutto il lavoro che le biblioteche hanno fatto in tutti questi decenni passati, nonché creare le condizioni effettive per quella cooperazione di cui nel settore bibliotecario si parla da tempo, ma che finora si è sviluppata in modo per certi versi limitato a causa delle difficoltà tecnologiche e delle resistenze psicologiche.

Pur non sottovalutando alcune questioni problematiche di natura tecnica e "filosofica" legate all'utilizzo di questa tecnologia e che devono essere affrontate criticamente, in questo contributo si è puntato a sottolineare i benefici legati alla possibilità di mettere in relazione risorse interne ed esterne alle varie istituzioni, garantendo dati autorevoli e di qualità utilizzabili da soggetti terzi i cui usi non sono prevedibili.

NOTE

¹ Proverbio africano.

² Cfr. MARIA TERESA BIAGETTI, *Sviluppi e trasformazioni delle biblioteche digitali: dai repositories di testi alle semantic digital libraries*, "AIB Studi", 54 (2014), n. 1, p. 11-34.

³ "L'inferenza è il processo deduttivo con il quale date due proposizioni si deduce la verità di una proposizione da un'altra giudicata vera, ovvero è la possibilità di dedurre conoscenza tramite percorsi. Ogni asserzione aggiunge nuova informazione, generando nuove inferenze e producendo nel Web una quantità sempre maggiore di informazione strutturata e classificata", ANTONELLA IACONO, *Linked data*, Roma, AIB, 2014.

⁴ "Con l'espressione linked data si fa dunque riferimento a dati pubblicati sul web in una modalità leggibile e interpretabile da una macchina, il cui significato sia esplicitamente definito tramite una stringa di parole e marcatori. Si costituisce così un reticolo di dati collegati appartenenti a un dominio (che costituisce il contesto di partenza), collegato a sua volta ad altri set di dati esterni, ovvero fuori dal dominio, in un contesto di relazioni sempre più complesso", MAURO GUERRINI - TIZIANA POSSEMATO, *Linked data: a new alphabet for the semantic web*, "JLIS.it", 4 (2013), n. 1, p. 67-97. Cfr. GIANFRANCO CRUPI, *Beyond the Pillars of Hercules: Linked data and Cultural Heritage*, "JLIS.it", 4 (2013), n. 1, p. 25-49.

⁵ CARLO BIANCHINI, *Library linked data e il futuro delle biblioteche*, <<https://sites.google.com/site/homepagecarlobianchini/Ricerca/library-linked-data-e-il-futuro-delle-biblioteche>>.

⁶ <<http://data.bnf.fr>>.

⁷ <<http://aims.fao.org/agrovoc>>.

⁸ <<http://www.w3.org/2004/02/skos/>>.

⁹ <<http://www.foaf-project.org>>.

¹⁰ ROMAIN WENZ, *Linked open data for new library services: the example of data.bnf.fr*, "JLIS.it", 4 (2013), n. 1, p. 403-415.

¹¹ <<http://creativecommons.org>>.

¹² <http://it.wikipedia.org/wiki/Application_programming_interface>.

¹³ <<https://www.ld4l.org>>.

¹⁴ "Currently, a lot of information about books, articles and cultural materials that could make them easier to find and understand is hidden in scattered systems across many libraries. We have an opportunity to use Linked Data as a common format to bring together all that scattered information. At Cornell, we've been working with the Semantic web for more than a decade, and we're very pleased that the Mellon Foundation has given us the opportunity to further that work with our colleagues at Harvard and Stanford", <<http://library.harvard.edu/12162013-1308/harvard-cornell-stanford-libraries-project-receives-grant>>.

¹⁵ "Our Lab is excited about the chance to work with such leading partners on a project that can make it easier for libraries to openly share more and more of what they know – not just about their content but about how that content is being used", <<http://library.harvard.edu/12162013-1308/harvard-cornell-stanford-libraries-project-receives-grant>>.

¹⁶ "One of the exciting prospects for this project is that we will be testing the usability and utility of the data that is produced. Many elements of this grant project overlap with the Technology Plan developed as part of the Stanford Linked Data Work-

shop in 2011 and provide us opportunity for great collaboration across three library systems”, <<http://library.harvard.edu/12162013-1308/harvard-cornell-stanford-libraries-project-receives-grant>>.

¹⁷ “The goal of the project is to create a Scholarly Resource Semantic Information Store (SRSIS) model that works both within individual institutions and through a coordinated, extensible network of Linked Open Data to capture the intellectual value that librarians and other domain experts and scholars add to information resources when they describe, annotate, organize, select, and use those resources, together with the social value evident from patterns of usage”. Cfr. <<https://www.ld4l.org>>.

¹⁸ <<http://www.vivoweb.org>>.

¹⁹ <<http://projecthydra.org>>.

²⁰ <<http://www.oclc.org/worldcat.en.html>>.

²¹ <<http://projecthydra.org>>.

²² <<http://linkeduniversities.org/index.html>>.

²³ <<http://data.open.ac.uk>>.

²⁴ <<http://www.w3.org/TR/void/>>.

²⁵ <<http://www.meducator.net>>.

²⁶ <<http://www.organic-edunet.eu/en>>.

²⁷ <<http://education.data.gov.uk>>.

²⁸ Cfr. <<http://lucero-project.info>> e MATHIEU D'AQUIN, *Linked Data for Open and Distance Learning* (second edition), <<http://www.col.org/PublicationDocuments/LinkedDataForODL.pdf>>. Al di fuori della rete Linked Universities è da segnalare il progetto anglossasone RLUNK, venti milioni di record bibliografici provenienti da trentaquattro biblioteche accademiche della Gran Bretagna e il tedesco HeBis, che mantiene ventisei milioni di metadati su ogni tipo di pubblicazione provenienti da 567 biblioteche accademiche in Asia, oltre al lavoro in corso a Oxford e Cambridge.

²⁹ ANTONELLA TROMBONE, *Da catalogare a identificare e collegare*, “Biblioteche oggi”, 32 (2014), n. 7, p. 4-8.

BIBLIOGRAFIA

MARISTELLA AGOSTI - NICOLA FERRO, *Interoperabilità tra sistemi di biblioteche digitali*, “DigItalia: rivista del digitale nei beni culturali”, 5 (2010), n. 1, p. 95-112.

MARIA TERESA BIAGETTI, *Sviluppi e trasformazioni delle biblioteche digitali: dai repositories di testi alle semantic digital libraries*, “AIB Studi”, 54 (2014), n. 1, p. 11-34; *Biblioteche digitali e semantic web*, “Bollettino AIB”, 49 (2009), n. 4, p. 577-583.

CARLO BIANCHINI - MAURO GUERRINI, *Introduzione a RDA. Linee*

guida per rappresentare e scoprire le risorse, Milano, Editrice Bibliografica, 2014.

CARLO BIANCHINI, *Library Linked Data e il futuro delle biblioteche*, <<https://sites.google.com/site/homepagecarlobianchini/Ricerca/library-linked-data-e-il-futuro-delle-biblioteche>>.

Bibliothèque nationale de France, <<http://data.bnf.fr>>.

GIANFRANCO CRUPI, *Beyond the Pillars of Hercules: Linked Data and Cultural Heritage*, “JLIS.it”, 4 (2013), n. 1, p. 25-49.

MATHIEU D'AQUIN, *Linked Data for Open and Distance Learning* (second edition), <<http://www.col.org/PublicationDocuments/LinkedDataForODL.pdf>>.

M. D'AQUIN et al., *Assessing the Educational Linked Data Landscape*, <<http://linkeduniv.eu/catalogue/publications/webosci2013-rn.pdf>>.

ANTONELLA DE ROBBIO, *Forme e gradi di apertura dei dati*, “Biblioteche oggi”, 30 (2012), n. 6, p. 11-24.

Global Interoperability and Linked Data in libraries (Firenze, 18-19 giugno 2012), “JLIS.it”, 4 (2013), n. 1.

MAURO GUERRINI - TIZIANA POSSEMATO, *Linked data: a new alphabet for the semantic web*, “JLIS.it”, 4 (2013), n. 1, p. 67-97.

ANTONELLA IACONO, *Linked data*, Roma, AIB, 2014; *Dal record al dato. Linked data e ricerca dell'informazione nell'OPAC*, “JLIS.it”, 5 (2014), n. 1, p. 77-102.

KATRINE KRIEGER - DIETMAR ROSNER, *Linked Data in e-learning: A Survey*, “SemanticWeb Journal”, n. 3, 2011.

Linked Data for Libraries, <<https://www.ld4l.org>>.

Linked universities, <<http://linkeduniversities.org/index.html>>.

FEDERICO MESCHINI, *Le mappe topiche. Come imparai a non preoccuparmi e ad amare i metadati*, “Bollettino AIB”, 45 (2005), n. 1, p. 59-72.

ALBERTO SALARELLI, *Sul perché, anche nel mondo dei Linked Data, non possiamo rinunciare al concetto di documenti*, “Studi AIB”, 54 (2014), n. 2-3, p. 279-293.

IVANA SARIĆ, *Metadata schemes significant for librarianship with the implementation of the OpenURL standard*, “Vjesnik bibliotekara Hrvatske”, 54 (2011), n. 1-2, p. 134-157.

ORESTE SIGNORE, *RDF per la rappresentazione della conoscenza*, <<http://www.w3c.it/papers/RDF.pdf>>.

EDMUNDO TOVAR - PIEDRA NELSON, *Linked Data to enhance users experience with Open Education Resources*, Cambridge, 16-28 aprile 2012, <http://oro.open.ac.uk/33640/5/Conference_Proceedings_Cambridge_2012.pdf>.

ANTONELLA TROMBONE, *Da catalogare a identificare e collegare*, “Biblioteche oggi”, 32 (2014), n. 7, p. 4-8.

ROMAIN WENZ, *Linked open data for new library services: the example of data.bnf.fr*, “JLIS.it”, 4 (2013), n. 1, p. 403-415.

DOI: 10.3302/0392-8586-201501-036-1

ABSTRACT

This article analyses the projects academic institutions and university libraries are launching to connect heterogeneous domains and resources by means of the tools of the Semantic web and the linked open data technology. In particular, three experiences are examined: the project by the Bibliothèque nationale de France, the Linked Data for Libraries project which involves the Cornell University Library, the Harvard Library Innovation Lab e the Stanford University Libraries, and the initiative carried on by some European universities participating in the Linked Universities network.