

# «Wikipedia loves libraries»: in Italia è un amore corrisposto...

di Anna Lucarelli

Sul fatto che le biblioteche stiano diventando sempre più un nodo fondamentale di una rete di relazioni che coinvolge altre realtà, altri attori e produttori di conoscenza, non ci sono dubbi.

Non è casuale il tema scelto per la Conferenza IFLA 2014 “Libraries, Citizens, Societies: Confluence for Knowledge”, ospitata a Lione nella Cité Internationale progettata da Renzo Piano proprio alla confluenza dei due grandi fiumi che attraversano la città, una confluenza che simboleggia energicamente l’incrocio di tradizioni, di visioni del mondo e di culture in cui sempre di più le biblioteche muovono i loro passi, orientano attività, intrecciano connessioni, sia contribuendo a produrre esse stesse conoscenza, sia iniziando a beneficiare delle varie forme di creatività collettiva e partecipativa favorite dal Web.

In questo panorama, una delle più interessanti novità si è sviluppata a partire dal 2011 nell’alveo del progetto GLAM<sup>1</sup> (sigla che combina le iniziali di *Galleries, Libraries, Archives, Museums*, istituzioni della memoria a cui si sono aggiunti di recente i giardini zoologici e botanici), un progetto avviato dalla comunità di Wikipedia, l’enciclopedia collaborativa online (ormai in 287 lingue) allestita da volontari e sostenuta da Wikimedia Foundation<sup>2</sup>, istituzione non profit che promuove un movimento mondiale di supporto allo sviluppo e alla condivisione libera e aperta dei contenuti della conoscenza, secondo una visione del sapere inteso come ‘bene comune’.

GLAM è nato dunque dalla consapevolezza, ormai consolidata, che queste istituzioni culturali, da sempre custodi di risorse, documentazione, immagini, catalogazioni, possono apportare un contributo fondamentale all’universo e alla comunità wikipediana e ha preso atto del ruolo innovativo che, in particolare le biblioteche, stanno assumendo grazie allo sviluppo dei *linked data*, i quali contri-

ANNA LUCARELLI, Biblioteca nazionale centrale Firenze, piazza Cavalleggeri 1/A, 50122 Firenze, e-mail [anna.lucarelli@beniculturali.it](mailto:anna.lucarelli@beniculturali.it).

Ultima consultazione siti web: 15 luglio 2014.

1 GLAM, <<https://outreach.wikimedia.org/wiki/GLAM>>.

2 Wikimedia, <<http://www.wikimedia.org/>>.



buiscono a dare un valore nuovo, più produttivo e aperto, ai contenuti e ai dati che esse producono<sup>3</sup>.

Così una delle linee di intervento di GLAM è diventata Wikipedia Loves Libraries, una campagna per potenziare il coinvolgimento e la collaborazione di Wikipedia con il mondo delle biblioteche, degli archivi e dell'open access<sup>4</sup>. Dalla prima Wikipedia Loves Libraries campaign (svoltasi nel 2011 in America del Nord) ad oggi, innumerevoli iniziative si sono realizzate in biblioteche di tutto il mondo – dalle più piccole e sconosciute alle più grandi e note – iniziative realizzate con l'obiettivo di mettere in comunicazione Wikipedia, straordinaria base di conoscenza con milioni di risorse informative, con il mondo delle biblioteche e di incrementare le voci dell'enciclopedia relative ad argomenti collegati ad esse o a qualsiasi altro tema del sapere<sup>5</sup>.

Come ha descritto Susanna Giaccai in un suo contributo sull'argomento ricco di informazioni, fra gli esempi più interessanti di attività che hanno coinvolto le biblioteche dobbiamo, ad esempio, menzionare:

Le sessioni di editazione collettiva organizzate dalla British Library per creare o migliorare la descrizione delle proprie collezioni in Wikipedia, oppure il progetto, più ampio e ben documentato, della State Library of Queensland, realizzato in collaborazione con Wikimedia Australia, volto a inserire contenuti di qualità relativi al proprio territorio coinvolgendo le biblioteche pubbliche locali. La collaborazione tra istituti della conoscenza e Wikipedia può anche realizzarsi con la donazione di immagini come nel caso degli Archives of American Arts che hanno donato a Wikimedia Commons (l'archivio di immagini libere) 256 immagini libere da diritti [...]<sup>6</sup>.

Contemporaneamente si è andata affermando la prassi dei wikipediani *in residence*. Istituzioni culturali e bibliotecarie di varia natura, quasi sempre molto prestigiose per la loro storia o per le funzioni che rivestono, hanno iniziato ad assumere persone che, come viene detto in una pagina appositamente dedicata a questo fenomeno:

dedicate time to working in-house at an organization. Wikipedians in Residence are usually financially compensated by the institution or by a Wikimedia chapter, but they may also be volunteers. The Wikipedian in Residence is not simply an in-house editor:

**3** Moltissimi i contributi su questo tema nella letteratura professionale: in Italia è stato trattato di recente anche da un esperto di tecnologia del Web semantico e sociale, di ingegneria della conoscenza e del progetto europeo LOD2, Sören Auer, *Il nuovo ruolo delle biblioteche in un mondo di scienza open e partecipativa, crowdsourcing e web data semantici*, in *La biblioteca connessa. Come cambiano le strategie di servizio al tempo dei social network*, *Convegno delle Stelline, Milano, 13-14 marzo 2014*, abstract, <<http://www.convegnostelline.it/relatori.php?IdUnivoco=87>>.

**4** Progetto GLAM biblioteche, <<http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Progetti>>. Una rassegna delle innumerevoli iniziative realizzate con istituzioni culturali di vario tipo, si trova nella *newsletter* mensile «This month in GLAM», <<http://outreach.wikimedia.org/wiki/GLAM/Newsletter>>.

**5** Wikipedia loves libraries, <[http://en.wikipedia.org/wiki/Wikipedia:Wikipedia\\_Loves\\_Libraries](http://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Loves_Libraries)>. Alla campagna è dedicato il portale Libraries. The Wikimedia way, <[http://outreach.wikimedia.org/wiki/Wikipedia\\_Loves\\_Libraries](http://outreach.wikimedia.org/wiki/Wikipedia_Loves_Libraries)>.

**6** Susanna Giaccai, *GLAM/Biblioteche: un esempio di collaborazione tra Wikipedia e biblioteche in Italia*, «Bibelot», (2013), n. 3-4, <<http://www.aib.it/struttura/sezioni/toscana/bibelot/bibelot-34-2013/2014/41014-glambiblioteche-un-esempio-di-collaborazione-tra-wikipedia-e-biblioteche-italia/>>.

the role is fundamentally about enabling the host organisation and its members to continue a productive relationship with the encyclopedia and its community after the Residency is finished. The Wikipedian in Residence model was first piloted by the GLAM initiative, but has since been adopted by other types of organizations[...]<sup>7</sup>.

Così, dal 2010 ad oggi, quasi un centinaio di persone, generalmente molto giovani, con le caratteristiche descritte e gli obiettivi indicati, sono state accolte da istituzioni di tutto il mondo: dal British Museum (che ha accolto il primo wikipediano *in residence* nel giugno del 2010) allo Château de Versailles; dagli US National Archives alla National Library of Israel; da OCLC (Online Computer Library Center) alla World Digital Library di Washington; dalla Biblioteca nazionale svizzera a quella spagnola. Ma sono moltissime le organizzazioni che si potrebbero citare e, fra le esperienze più recenti, anche quelle italiane attivate tra aprile e maggio del 2014, dalla biblioteca civica di Mestre e dal museo d'arte moderna e contemporanea di Trento e Rovereto (Mart). Insomma, quanto accaduto negli ultimi anni sembra proprio confermare un trend in ascesa, come se davvero quelle *7 reasons librarians should edit Wikipedia*, elencate recentemente da Natalie Binder, bibliotecaria e blogger statunitense, stiano diventando asserzioni sempre più condivise<sup>8</sup>.

In Italia, le prime attività legate al progetto GLAM che hanno attratto le biblioteche nell'universo wikipediano sono iniziate circa due anni fa quando alcuni bibliotecari italiani come Andrea Zanni (oggi presidente di Wikimedia Italia), Susanna Giaccai, Virginia Gentilini, e altri appassionati di Wikimedia Italia come Cristian Consonni e Pierfranco Minsenti, hanno cominciato a lavorare per favorirne la diffusione. Proprio a Zanni, Giaccai e Gentilini dobbiamo le prime uscite sull'argomento grazie ai loro contributi pubblicati a partire dal 2011<sup>9</sup>, in verità affiancati da un'intervi-

**7** Wikipedian in Residence, <[http://outreach.wikimedia.org/wiki/Wikipedian\\_in\\_Residence](http://outreach.wikimedia.org/wiki/Wikipedian_in_Residence)>. Le *core characteristics* di un wikipediano *in residence*: «Serves as a liaison between the organization and the Wikimedia community to promote a mutually beneficial cooperation. Promotes understanding of Wikimedia among staff at the organization through workshops and events. Works with organizational staff to digitize, compile, and organize resources that can be shared with the Wikipedia community. Facilitates the improvement of content by the Wikipedia community, rather than directly editing articles as a core goal. Coordinates events, such as Hack-a-Thons, Edit-a-Thons, or Backstage Passes, that bring Wikipedians on-site to work with staff on content creation and improvement. Minimizes impacts of Conflict of Interest by not editing articles directly relating to the organization. Formally coordinated by the institution, allowing the resident to work closely with staff for various projects. Ultimately, a residency lays the foundation for a more lasting partnership between the Wikipedia community and an organization».

**8** Natalie V. Binder, *Information is hierarchy*, <<http://nrbinder.wordpress.com/2014/03/22/librarians-should-edit-wikipedia/>>.

**9** Mi permetto di indicarne alcuni: Andrea Zanni, *Le biblioteche e l'universo Wikimedia*, «AIB Notizie», 24 (2012), n. 2, <<http://www.aib.it/pubblicazioni-aib/aib-notizie/2012/25605-biblioteche-e-wikimedia/>>; Susanna Giaccai, *La biblioteca in Wikipedia*, «AIB Notizie», 24 (2012), n. 3, <<http://www.aib.it/pubblicazioni-aib/aib-notizie/2012/28267-biblioteca-in-wikipedia/>>; Virginia Gentilini, *Librarians are Wikipedians Too. La collaborazione tra biblioteche e progetti wiki, vista dall'Italia*, «Bibliotime», 15 (2012), n. 3, <<http://www.aib.it/aib/sezioni/emr/bibtime/num-xv-3/gentilini.htm>>. Per una selezione esaustiva di articoli e contributi italiani vedere la pagina *Progetto:GLAM/Biblioteche/Lettere consigliate*, <[http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Lettere\\_consigliate](http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Lettere_consigliate)>.

sta di Zanni a Umberto Eco, effettuata il 24 aprile del 2010 e pubblicata fra il 2010 e il 2012, che in qualche modo ne anticipava alcuni temi<sup>10</sup>.



**Figura 1 - Il logo della versione in italiano di Wikipedia**

Rinviando al già citato articolo di Giaccai sui primi passi di questo percorso, ritengo comunque utile sottolineare sinteticamente alcuni aspetti particolarmente interessanti. Uno di questi è la velocità con cui, in un lasso di tempo brevissimo, anche in Italia sono stati realizzati eventi e organizzate iniziative, seminari tecnici, conferenze, corsi di formazione per bibliotecari o utenti delle biblioteche, introduttivi a Wikipedia oppure dedicati alla redazione delle sue voci<sup>11</sup>. Ma nell'ultimo anno sono state anche siglate convenzioni ed elaborati veri e propri progetti, come nel caso di quanto si sta realizzando alla Biblioteca nazionale centrale di Firenze.

A favorire l'incontro delle biblioteche italiane con l'universo wikimediano sono intervenuti vari fattori, alcuni dei quali hanno fatto leva su caratteristiche, non a tutti conosciute, della comunità wikipediana. Questa, nonostante sia ispirata a principi di collaborazione libera e aperta, contrariamente a quanto molti immaginavano fino a poco tempo fa, si muove in un contesto di regole (auto-regole) e metodologie definite, condivise e per certi versi rigorose, alle quali per ogni *contributor* è d'obbligo adeguarsi, come si può facilmente dedurre dalle istruzioni del Portale comunità<sup>12</sup> dedicato a politiche, linee guida e specifiche guide, manuali operativi, informazioni tecniche ecc.

**10** Andrea Zanni, *Intervista a Umberto Eco*, <[http://it.wikinews.org/wiki/Intervista\\_a\\_Umberto\\_Eco](http://it.wikinews.org/wiki/Intervista_a_Umberto_Eco)>; *Wikimedia Italia intervista Umberto Eco (I parte)*, «AIB notizie», 22 (2010), n. 5, <<http://www.aib.it/aib/editoria/n22/0510.htm3>>; *Wikimedia Italia: intervista a Umberto Eco (seconda parte)*, «AIB notizie», 22 (2010), n. 6, <<http://www.aib.it/aib/editoria/n22/0610.htm3>>; *Wikimedia Italia: intervista a Umberto Eco (terza parte)*, «AIB notizie», 23 (2011), n. 1, <<http://www.aib.it/aib/editoria/n23/0103.htm3>>.

**11** Elenco delle iniziative svolte dal 2013 ad oggi, <<http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Progetti>>.

**12** Portale:Comunità, <<http://it.wikipedia.org/wiki/Portale:Comunit%C3%A0>>.

Dunque, non è affatto un caso che uno dei primi ambiti di convergenza sia avvenuto in relazione ad attività di formalizzazione e normalizzazione di dati di tipo catalografico e sia stato sollecitato, anche in Italia, dalla diffusione di informazioni su un'iniziativa di OCLC, il già citato consorzio di biblioteche nato alla fine degli anni Sessanta che, tramite *WorldCat*, mette a disposizione descrizioni bibliografiche fornite da cataloghi nazionali e internazionali di vari paesi del mondo. OCLC, proprio grazie ad un wikipediano *in residence* Max Klein, aveva studiato come potenziare la fruizione delle biblioteche da parte degli utenti del Web, avviando una condivisione di dati fra la versione inglese di Wikipedia e il VIAF (Virtual International Authority File), che proprio OCLC gestisce, per raccordare e mettere a disposizione milioni di voci di autorità riferite a nomi propri di persone, organizzazioni, titoli di opere e luoghi<sup>13</sup>. Come è ormai noto, anche grazie al lavoro pionieristico di Klein, altre versioni linguistiche di Wikipedia si sono via via avvicinate alle biblioteche, mettendo a disposizione del loro vasto pubblico e, in particolare dei lettori delle voci di tipo biografico, dati controllati e di qualità che consentono di raccordare forme varianti, pseudonimi ecc.; si tratta in realtà di un'opportunità reciproca, dato che anche le biblioteche, tramite questi collegamenti, possono diffondere il risultato del proprio lavoro in contesti di open access e al di fuori del veicolo tradizionale e 'concluso' dei cataloghi. E non va dimenticato che anche altri *data set* prodotti dalle biblioteche come, ad esempio, gli schemi di classificazione, potranno essere progressivamente implementati per *interlinking* di diversa natura.

Il coinvolgimento delle biblioteche italiane agli sviluppi della collaborazione fra VIAF e Wikipedia è stato allora uno dei temi su cui si è maggiormente mosso, sin dall'inizio, l'interesse di wikipediani e bibliotecari. Si è trattato di un'utilissima occasione per analizzare e far conoscere l'attività che l'ICCU, in quanto collettore dei record di autorità prodotti nell'ambito del Servizio bibliotecario nazionale (SBN) secondo il formato UNIMARC/A, svolge come partner del progetto VIAF dal 2009<sup>14</sup>, ma anche per portare alla luce le carenze e le esiguità dell'acquisizione dei record italiani, le difficoltà di *matching* e altre problematiche che hanno comportato sinora una ridotta visibilità delle notizie bibliografiche di SBN e dunque una partecipazione dell'Italia più limitata rispetto a quella di altri paesi<sup>15</sup>.

Il lavoro di analisi svolto su queste questioni ha fornito materiali preziosi sia per il convegno dal titolo "Rilanciare il Servizio Bibliotecario Nazionale (SBN)", svolto a Roma il 20 giugno del 2013<sup>16</sup>, sia per la stesura di un documento dell'Associazione

**13** Per informazioni tecniche sul progetto: Maximilian Klein; Alex Kyrios, *VIAFbot and the Integration of Library Data on Wikipedia*, «Code4Lib Journal», (2013), n. 22, <<http://journal.code4lib.org/articles/8964>>. Interessante anche il video: Max Klein, *Wikidata and Authority Control Video Explains How Wikidata Bolstered Authority Control in Wikipedia*, <<http://www.oclc.org/research/news/2013/06-13.html>>.

**14** Istituto Centrale per il Catalogo Unico, *Virtual International Authority File (VIAF)*, <[http://www.iccu.sbn.it/opencms/opencms/it/main/attivita/internaz/pagina\\_367.html](http://www.iccu.sbn.it/opencms/opencms/it/main/attivita/internaz/pagina_367.html)>.

**15** Una presentazione approfondita e documentata del progetto VIAF in relazione a Wikipedia e delle problematiche legate alla partecipazione italiana è stata presentata da Pierfranco Minsenti, *Un'altra modalità per arricchire Wikipedia: l'esperienza dell'importazione dei dati dal Virtual International Authority File*, in *BiblioPride. Giornata nazionale delle biblioteche. Biblioteche e Wikipedia: condivisione, open data e competenze*, Firenze, Biblioteca Marucelliana, 5 ottobre 2013.

**16** Un resoconto del convegno si trova in Lucia Antonelli, *Rilanciare SBN: descrizione di un'opera in continuazione*, «AIB Notizie», 25 (2013), n. 4, <<http://www.aib.it/attivita/2013/37874-rilanciare-sbn/>>.

ne italiana biblioteche<sup>17</sup> al quale l'ICCU ha risposto con una dichiarazione ufficiale di impegni fra cui, come indicato nel paragrafo *Finalità e governance del servizio*, quello di aprire sempre di più SBN «alla collaborazione ed interscambio dei dati con il mondo dell'editoria, degli archivi, dei musei e, non ultima, di Wikipedia». Nella risposta dell'ICCU si legge anche che «è stato attivato il link che da Wikipedia rimanda alle schede di autorità presenti nell'OPAC SBN. È stato prodotto un *permalink* per permettere un più agevole puntamento alle notizie analitiche dei documenti»; e più oltre, nel paragrafo su *Catalogazione e manutenzione del catalogo*, dichiara di ritenere fondamentale «il proseguimento e l'intensificazione della partecipazione dell'Italia al VIAF», specificando che verranno ad esso inviati «autori a livello 95 creati dalle Biblioteche nazionali di Firenze e Roma (circa 140.000), anche se parte di questi non hanno una scheda d'autorità completa»<sup>18</sup>.

Nello specifico, su Wikidata è stata creata il 6 aprile 2013 una proprietà<sup>19</sup> che permette di collegare gli elementi (*item*) di Wikidata alle schede di autorità del Servizio bibliotecario nazionale, tramite il *permalink* citato dal documento ICCU. Al momento, la proprietà P396 (identificativo SBN) connette 2.530 schede di autorità ai corrispettivi elementi su Wikidata, il che rende possibile richiamare il collegamento delle relative voci su Wikipedia. Inoltre, l'ICCU da gennaio 2014 può beneficiare della collaborazione di un wikipediano, Luca Martinelli (attuale segretario dell'associazione Wikimedia Italia), che è stato incluso nel gruppo di lavoro per la pubblicazione dei *data set* riguardanti il materiale bibliografico dell'OPAC relativo ai fondi moderno e antico. Al momento, è stata conclusa la prima fase del lavoro iniziata a gennaio 2014 e consistente nella mappatura con il modello FRBRoo<sup>20</sup>, secondo il modello concettuale di riferimento CIDOC-CRM, degli elementi UNIMARC considerati rilevanti per l'esportazione in *linked open data*. Lo scopo sarà la pubblicazione di tali dati in un formato interoperabile, per permettere la loro massima diffusione anche sui progetti Wikimedia, in particolare Wikidata<sup>21</sup>.

In Italia l'attenzione delle biblioteche e dei bibliotecari per Wikipedia si è sviluppata contemporaneamente su più fronti, facilitata anche dalla creazione di una lista di discussione ([bibliotecari@wikimedia.it](mailto:bibliotecari@wikimedia.it)), nata nel settembre del 2012 e inizialmente frequentata da pochi appassionati, ma allargatasi in pochissimo tempo a simpatizzanti di varie professionalità e competenze. Si tratta di una lista in cui i bibliotecari italiani hanno iniziato a socializzare e confrontarsi con altri tipi di comunità, processo da cui sono derivati effetti e risultati che si sono spinti velocemente molto più in là di quanto si potesse immaginare.

**17** Associazione Italiana Biblioteche, *AIB per il #nuovoSBN. Documento approvato dal CEN il 17 dicembre 2013*, <<http://www.aib.it/wp-content/uploads/2013/12/nuovoSBN.pdf>>.

**18** Testo completo della *Risposta ICCU al documento AIB per il #nuovoSBN*, <<http://www.aib.it/attivita/2014/41208-iccu-nuovosbn/>>.

**19** Una proprietà su Wikidata è l'equivalente di un campo in un database, ossia è la descrizione semantica di un valore.

**20** FRBRoo è un'ontologia volta a rappresentare la semantica alla base delle informazioni bibliografiche e a facilitare l'integrazione e lo scambio con il mondo museale, già utilizzata dalla Bibliothèque nationale de France per la pubblicazione dei propri *data set*, <[http://www.cidoc-crm.org/frbr\\_inro.html](http://www.cidoc-crm.org/frbr_inro.html)>.

**21** Un ringraziamento all'ICCU per le informazioni su questo work in progress.

Come ha raccontato Susanna Giaccai, tra i promotori della lista:

il mio obiettivo era di incrementare la presenza delle biblioteche in Wikipedia creando occasione di formazione sulle norme editoriali di Wikipedia per i bibliotecari e gli utenti delle biblioteche, quello di Andrea Zanni, interessato alle problematiche di riuso degli *open data*, di acquisire dalle biblioteche dati di qualità per Wikipedia e quello di Virginia Gentilini di analizzare le possibilità offerte da Wikipedia nell'esplorare i nuovi compiti della biblioteca. Sul tema della collaborazione tra biblioteche e Wikipedia, in quelle settimane, erano usciti alcuni scritti su bollettini professionali e in blog ed erano passate alcune mail in AIB-CUR. Il gruppetto attorno alla lista [bibliotecari@wikimedia.it](mailto:bibliotecari@wikimedia.it) si andava così allargando attraverso il coinvolgimento di altri wikipediani e bibliotecari. Ad agosto Stefano Parise mi dichiarò l'interesse dell'AIB per l'iniziativa; in lista all'inizio si discusse di come creare una versione italiana della pagine del Progetto GLAM anglofono<sup>22</sup>.

Di lì a poco, a gennaio del 2013, anche chi scrive intervenne nella lista, segnalando non solo l'interesse della Biblioteca nazionale centrale di Firenze ad inserirsi nel percorso già intrapreso da biblioteche nazionali di altri paesi per rendere sempre più disponibili in forma aperta i metadati catalografici da esse prodotte, ma in particolare per informare della disponibilità ad approfondire modalità di implementazione di collegamenti fra Wikipedia e il thesaurus del *Nuovo soggettario*, lo strumento italiano di indicizzazione per soggetto in forma di vocabolario controllato, allestito, sviluppato e aggiornato dalla BNCf e disponibile in formati adeguati al Web semantico<sup>23</sup>.

L'attivazione di forme di *linking* fra strumenti catalografici italiani e Wikipedia sono favorite dalle funzionalità della DBpedia Italiana<sup>24</sup> il cui database, ricco di milioni di dati aperti in formato RDF, è realizzato dalla Fondazione Bruno Kessler di Trento e che, come dichiarato, consente «l'estrazione e il riutilizzo di informazioni semanticamente strutturate dalla versione italiana di Wikipedia. Il progetto mira a rendere riutilizzabili le informazioni di Wikipedia da parte di software e applicazioni. La DBpedia Italiana permette di eseguire *query* sofisticate e complesse sui contenuti di Wikipedia e di collegare altri *data set Linked Data* a Wikipedia».

La ricchezza dei 'dati qualificati' gestiti dalla versione inglese di DBpedia ha fatto sì che il progetto si collocasse al centro della *Linked Open Data cloud* e che nel 2012, su stimolo di Wikimedia Deutschland, fosse avviato anche Wikidata, una base di conoscenza che contiene dati in tutte le lingue in cui Wikimedia Foundation ha attivato i suoi progetti e che costituisce «un accesso centralizzato alla gestione di dati strutturati» provenienti da varie fonti<sup>25</sup>.

<sup>22</sup> Susanna Giaccai, *GLAM/Biblioteche: un esempio di collaborazione tra Wikipedia e biblioteche in Italia*, cit.

<sup>23</sup> Sulla strada intrapresa dalla Nazionale di Firenze verso i *linked data*, viene presentato il contributo: Anna Lucarelli, *Esperienze di linked data e web semantico in BNCf*, in *6. Incontro ISKO Italia, Firenze, 20 maggio 2013*, <<http://www.iskoi.org/doc/firenze13/lucarelli.pdf>>.

<sup>24</sup> DBpedia Italiana, <<http://it.dbpedia.org/>>.

<sup>25</sup> Wikidata, <[http://www.wikidata.org/wiki/Wikidata:Main\\_Page](http://www.wikidata.org/wiki/Wikidata:Main_Page)>. Su Wikidata e le sue potenzialità, recentemente anche Luca Martinelli, *Wikidata: A New Way to Disseminate Structured Data*, in *FSR 2014. Faster, smarter and richer. Reshaping the library catalogue. International conference, Roma, 27-28 febbraio 2014*, <<http://www.slideshare.net/SannitaSSJ/presentazione-aib-martinelli>>.

### La Biblioteca nazionale centrale di Firenze e la collaborazione con Wikipedia

Così, in un lasso di tempo molto breve, si è realizzato il primo prototipo italiano di convergenza fra Wikipedia e uno strumento catalografico prodotto da una biblioteca; i risultati del lavoro sono stati presentati, a distanza di pochi mesi dall'avvio, nel mese di ottobre 2013, durante l'edizione fiorentina del già citato *Bibliopride*, in un seminario dedicato, appunto, a "Biblioteche e Wikipedia: condivisione open data e competenze"<sup>26</sup>. Il seminario prevedeva come relatori bibliotecari, wikipediani e wikimediani e fra loro anche Frieda Brioschi, nel ruolo, rivestito in quel momento, di presidente di Wikimedia Italia. Per la Biblioteca nazionale centrale di Firenze è stata l'occasione per presentare quanto realizzato e per raccontare un percorso impostato in realtà anni addietro, quando la Biblioteca aveva iniziato ad occuparsi di Wikipedia in occasione dei primi lavori sul *Nuovo soggettario*, ed era successivamente proseguita nell'ambito degli sviluppi che lo strumento ha avuto sul fronte dei *linked data*<sup>27</sup>.

Come è noto, la componente principale di questo sistema di indicizzazione semantica è un thesaurus multidisciplinare, disponibile online, in continuo ampliamento. Attualmente contiene oltre 50.000 termini e si alimenta anche grazie alle proposte di altre istituzioni che collaborano al progetto. Il thesaurus non si limita ad essere un 'attrezzo di lavoro' di chi si occupa di catalogazione semantica. Oltre al collegamento con gli OPAC ad esso integrabili, ha altre funzionalità che lo rendono utilizzabile per ottimizzare il recupero dell'informazione, ampliare i percorsi della ricerca, navigare verso altre risorse online. Per saperne di più, si può leggere una descrizione nella pagina web che gli dedica la BNCF, oppure alla voce *Nuovo soggettario* di Wikipedia<sup>28</sup>. La disponibilità dei suoi dati nel formato standard SKOS (Simple Knowledge Organization System) ne rende possibile l'interoperabilità con altri servizi, la fruibilità nel mondo dei *linked data* e del Web semantico: i metadati sono esposti in forma di triple RDF e dunque sono riusabili anche in contesti non necessariamente bibliotecari<sup>29</sup>.

L'interesse della BNCF per Wikipedia in lingua italiana dunque è iniziato nel 2005, agli esordi dei lavori del prototipo del thesaurus, quando sull'enciclopedia libera e collaborativa, come essa stessa si definisce, circolava ancora scetticismo, perché ritenuta poco affidabile in quanto non controllata e priva di una redazione che ne garantisce la qualità dei contributi.

Sulle analogie e le diversità che legano thesauri ed enciclopedie si è già riflettuto in altre sedi; non hanno natura e finalità omologabili, ma sono entrambi strumenti di organizzazione del sapere e di approfondimento della conoscenza, sono strutturati su basi logiche e/o categoriali, sono basati sull'uso del linguaggio e delle

<sup>26</sup> Biblioteche e Wikipedia: condivisione open data e competenze, <[http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Progetti/Biblioteche\\_e\\_Wikipedia\\_condivisione\\_open\\_data\\_e\\_competenze](http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Progetti/Biblioteche_e_Wikipedia_condivisione_open_data_e_competenze)>.

<sup>27</sup> Giovanni Bergamin; Anna Lucarelli, *Il progetto di interoperabilità del Thesaurus BNCF con Wikipedia*, in *Bibliopride. Giornata nazionale delle biblioteche. Biblioteche e Wikipedia: condivisione, open data e competenze, Firenze, Biblioteca Marucelliana, 5 ottobre 2013*, <<http://www.slideshare.net/SusannaG/bibliopride-2013-slides-lucarelli>>.

<sup>28</sup> *Nuovo soggettario*, <<http://thes.bncf.firenze.sbn.it/>>; <[http://it.wikipedia.org/wiki/Nuovo\\_soggettario](http://it.wikipedia.org/wiki/Nuovo_soggettario)>.

<sup>29</sup> SKOS, come si può leggere nella home page, è definito «an area of work developing specifications and standards to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading lists and taxonomies within the framework of the Semantic Web», <<http://www.w3.org/2004/02/skos/>>.

parole come significanti polifunzionali. E d'altra parte, sappiamo quanto il Web abbia contribuito ad avvicinare sempre più strutture, funzionalità e caratteristiche di apparati lessicografici ed enciclopedici, di motori di ricerca e di strumenti di indicizzazione semantica<sup>30</sup>.

Non è un caso che, già qualche anno fa, si siano avviati progetti sulla possibilità di creare metadati semantici sfruttando le potenzialità di Wikipedia, come nel caso di Concept Mapper, uno strumento che il Cineca ha sviluppato all'interno del progetto europeo Papyrus, con l'obiettivo di sperimentare l'indicizzazione automatica di risorse testuali di ambito tecnico e scientifico tramite identificazione di concetti nel contesto dei documenti, eventuali link a pagine dell'enciclopedia, associazione di metadati semantici e mappatura con specifiche ontologie<sup>31</sup>.

Dunque, nonostante un thesaurus non intenda essere né assomigliare ad un'enciclopedia, la natura aperta di Wikipedia la rendeva particolarmente attraente per il gruppo di lavoro del *Nuovo soggettario*, che ne teneva d'occhio senza pregiudizi gli sviluppi, pur usando contemporaneamente come strumenti di lavoro e fonti per ricerche lessicografiche e semantiche anche altre enciclopedie, tradizionalmente considerate molto più prestigiose (come la Britannica o la Treccani), le quali proprio in quegli anni rendevano progressivamente disponibili (completamente o parzialmente) edizioni online dei loro contenuti<sup>32</sup>.

Inizialmente, la redazione del *Nuovo soggettario* usava Wikipedia come uno degli strumenti di riferimento per controlli morfologici o semantici relativi a termini che venivano via via accolti nel thesaurus. L'enciclopedia veniva citata nel campo 'Fonte' dei termini del thesaurus a cui corrispondevano lemmi nella stessa forma. In quei primi anni, insomma, era considerata una fonte di ricerche come altre, peraltro di natura residuale, perché veniva consultata (e citata) solo nei casi in cui non era possibile rintracciare attestazioni in opere considerate scientificamente più attendibili. Si trattava di un interesse che risentiva di una logica tradizionale, di una visione 'referenziale' ancora lontana dal mondo dei *linked open data* a cui soltanto qualche anno dopo sia Wikipedia che il *Nuovo soggettario* si sarebbero affacciati.

Nel 2005 di Web semantico (il 'Web dei dati' nella ormai classica definizione di Tim Berners-Lee) non si parlava ancora e, nonostante la BNCF avesse già espresso

**30** Anna Lucarelli, *I destini incrociati di strumenti di indicizzazione, dizionari, enciclopedie, motori di ricerca*, «Biblioteche oggi», (2012), n. 9, p. 5-14, <<http://www.bibliotecheoggi.it/content/201200900501.pdf>>. Sui rapporti fra diverse tipologie di sistemi di organizzazione della conoscenza, recentemente anche Claudio Gnoli, *Connettere la conoscenza: quattro livelli di interoperabilità*, in *La biblioteca connessa. Come cambiano le strategie di servizio al tempo dei social network*, *Convegno delle Stelline*, Milano, 13-14 marzo 2014.

**31** Giulio Paci; Giorgio Pedrazzi; Roberta Tura, *Concept Mapper. Da Wikipedia ai metadati semantici*, «Notizie dal Cineca», (2011), n. 66, p. 25-28, <<http://www.cineca.it/sites/default/files/cineca/publicazioni/scaricabili/Notiziario66.pdf>>.

**32** Il dibattito su logiche, caratteristiche e funzionalità delle enciclopedie digitali in relazione a quelle cartacee è molto fervido sia a livello internazionale che nazionale. Interessanti le opere come *La revolution Wikipedia. Les encyclopedies vont-elles mourir?*, prefate de Pierre Assouline. [Paris]: Mille et une nuits, c2007; oppure, anche in edizione italiana, Peter Burke, *Dall'Encyclopédie a Wikipedia. Storia sociale della conoscenza*. Bologna: Il mulino, 2013. Il portale Treccani ha pubblicato recentemente in video un'intervista di Massimo Bray sull'argomento: *Tullio Gregory: Il futuro delle enciclopedie. Parte I-II-III*, <[http://www.treccani.it/webtv/videos/Int\\_Tullio\\_Gregory\\_il\\_futuro\\_delle\\_enciclopedie\\_parte\\_I.html](http://www.treccani.it/webtv/videos/Int_Tullio_Gregory_il_futuro_delle_enciclopedie_parte_I.html)>.

l'intenzione di adottare per il thesaurus protocolli orientati al Web per renderlo fruibile da altre applicazioni, la sua pubblicazione era ancora limitata al formato Zthes<sup>33</sup>, in quel momento il modello internazionalmente più condiviso per la rappresentazione e la navigazione di vocabolari controllati.

Il primo termine del *Nuovo soggettario* a riportare nel campo 'Fonte' la citazione di Wikipedia viene inserito nel luglio del 2005: si tratta di *RFID* <*Radio frequency identification*>, non a caso un termine 'non preferito' che rinvia a Identificazione a radiofrequenza. Ma si trattava, appunto, soltanto di una citazione, statica e senza alcuna forma di collegamento.

Un grande cambiamento è avvenuto nel 2007, quando la dicitura Wikipedia che corredeva come 'Fonte' alcuni termini ha iniziato ad indicare non soltanto uno strumento di reference ma un vero e proprio link, inizialmente puntato alla home page di Wikipedia e successivamente, a distanza di un mese, collegato specificamente alla pagina del lemma corrispondente<sup>34</sup>. Da questo momento, Wikipedia non è stata più soltanto una fonte di attestazione ma un vero e proprio collegamento esterno.

The screenshot shows the 'Nuovo soggettario THESAURUS' interface from the Biblioteca Nazionale Centrale Firenze. The main content area displays the entry for 'Frutteti'. The 'Fonte' field contains the text: 'Soggettario; Treccani.it; DeM; NDTA; Wikipedia(IT)'. Below this, other fields are listed: 'Classificazione Dewey (Ed. 23): 634', 'Equiv. LCSH: Orchards', 'Agenzia catalografica/Proponente: BNI; Biblioteca Bocconi', 'Status del record: Termine strutturato', and 'Identificativo: 41805'. The interface also includes a search sidebar on the left and a navigation menu at the top.

Figura 2 – Termine del *Nuovo soggettario* con link (dal campo Fonte) a Wikipedia.

Questo percorso di apertura e interoperabilità, sia semantica che tecnica, del thesaurus verso altri sistemi di organizzazione e diffusione dalla conoscenza si era in verità allargato anche ad altri strumenti. In parallelo, si incrementavano collega-

<sup>33</sup> The Zthes specifications for thesaurus representation, access and navigation, <<http://zthes.z3950.org/>>.

<sup>34</sup> Può essere interessante segnalare che il primo termine del *Nuovo soggettario* che, attraverso la citazione Wikipedia nel campo 'Fonte', ha attivato il collegamento alla home page dell'enciclopedia è *Imposta negativa sul reddito* (aprile 2007), mentre il primo che ha attivato un *deep link* al corrispondente lemma è *Economia industriale* (maggio 2007).

menti con altri thesauri (ad esempio con Agrovoc, con ThESS il thesaurus di economia e scienze sociali della Università di Castellanza, con Eurovoc ecc.), con i lemmi della Treccani online, con i soggetti in inglese della *Library of Congress Subject Headings*<sup>35</sup>, tutti collegamenti che nel tempo sarebbero ulteriormente esplosi, potendo beneficiare, in alcuni casi, di comuni formati di marcatura e principalmente del già citato SKOS/RDF<sup>36</sup>.

Un'ulteriore e più evoluta fase dei rapporti fra il *Nuovo soggettario* e Wikipedia è iniziata nel 2010, attraverso un percorso favorito da fattori di natura non soltanto biblioteconomica. Un ruolo determinante è stato svolto dal progressivo e costante superamento della percezione sociale e culturale di Wikipedia come strumento di bassa qualità, un cambiamento lentamente ma efficacemente affermatosi a più livelli, basti pensare alla versione online della Treccani, simbolo di un prodotto editoriale di alto livello, che da qualche anno, e in certi specifici casi, fornisce citazioni e collegamenti a Wikipedia. Ma a facilitare e rendere più stretto l'incontro con il *Nuovo soggettario* hanno senz'altro influito anche altri fattori di natura tecnica. Mentre il thesaurus intraprendeva un percorso di apertura verso altri strumenti di organizzazione della conoscenza nella logica dei *linked data*, come è noto, DBpedia assumeva, come già accennato, il ruolo di nucleo centrale della *Linked Data Cloud*<sup>37</sup>.

Ma è interessante sottolineare che il progressivo legame fra *Nuovo soggettario* e Wikipedia sia stato facilitato anche da fattori di natura *social* e partecipativa, come il coinvolgimento dei bibliotecari della BNCF che si occupano di indicizzazione semantica alla già citata [bibliotecari@wikimedia.it](mailto:bibliotecari@wikimedia.it), una lista cresciuta in fretta e che ha avuto un ruolo propulsivo e stimolante poiché ha messo in contatto persone appartenenti a mondi diversi, che hanno avanzato proposte e iniziative concrete di collaborazione. La lista ha favorito i primi contatti diretti fra bibliotecari della BNCF e wikimediani e ha consentito l'organizzazione di una serie di incontri che, da marzo del 2013, sono diventati anche 'non virtuali' e sono proseguiti sino ad oggi.

Questo approccio fra bibliotecari (non sempre giovani) e wikipediani (quasi sempre giovani e abili anche sul fronte informatico), inizialmente un po' sospettoso e diffidente, si è velocemente trasformato in un'esperienza stimolante e fruttuosa. Così i collegamenti fra il thesaurus e Wikipedia, supportati dal linguaggio RDF impiegato in entrambi i *data set*, hanno iniziato ad accrescersi e a perfezionarsi qualitativamente, mettendo in risalto la necessità di affrontare problematiche di interoperabilità sia tecnica che semantica legate alla polisemia, a criteri di disambiguazione, all'analisi categoriale ecc. Era ovvio che non tutti i criteri seguiti per allestire il thesaurus potessero, né dovessero, coincidere con quelli usati dai *contributors* di Wikipedia ed era ben chiaro che nessuno dei due strumenti dovesse imporsi sulle logiche dell'altro. Non avrebbe avuto senso, ad esempio, costringere i wikipediani a seguire i principi di analisi categoriale impiegati per il thesaurus, oppure adottare i criteri di preferenza del singolare o plurale imposti dagli standard sui vocabolari controllati. Al contrario, altre funzionalità avrebbero potuto avere un'utilità condivisa come, ad esempio, l'apparato di relazioni di equivalenza (sinonimiche e quasi sino-

**35** Library of Congress Subject Headings, <<http://id.loc.gov/authorities/subjects.html>>.

**36** Ai dati quantitativi specifici (aggiornati al marzo 2014) sui collegamenti esterni del thesaurus con altri sistemi di organizzazione della conoscenza è dedicata una pagina nel sito della BNCF, <<http://digitale.bncf.firenze.sbn.it/NS-SPARQL/>>.

**37** Linked Data. Connect Distributed Data across the Web, <<http://linkeddata.org/>>.

nimiche) che il thesaurus metteva a disposizione della comunità dei collaboratori e lettori dell'enciclopedia, oppure delle relazioni associative, un tipo di collegamento semantico che l'enciclopedia avrebbe in effetti da qualche tempo iniziato ad adottare tramite i link alle voci correlate, inseriti a fondo pagina di molti lemmi.

Il vero salto qualitativo e quantitativo è avvenuto nei primi mesi del 2013, quando gli informatici della BNCF hanno messo a punto una sorta di algoritmo che ha consentito di sperimentare l'acquisizione automatica di link a Wikipedia, senza che i curatori del *Nuovo soggettario* dovessero inserirli manualmente e a seguito di un'operazione di tipo intellettuale. I collegamenti fra i termini del thesaurus e i corrispondenti lemmi dell'enciclopedia (realizzati tramite il formato RDF/XML disponibile sulla DBpedia italiana e accessibile attraverso un *endpoint* SPARQL) si sono così vertiginosamente sviluppati. Soltanto in un anno, fra ottobre del 2012 e ottobre del 2013, sono passati da 990 a 12.350, raggiungendo ad aprile 2014 quota 13.530 (cifra che comprende i collegamenti a pagine di disambiguazione) e costituendo il maggior numero di collegamenti esterni del *Nuovo soggettario*, seguiti numericamente da quelli con la Treccani online (circa 9.500) e da quelli con LCSH (*Library of Congress Subject Headings*), implementati ad oggi per oltre 5.200 descrittori. Altre modalità di *matching* sono in via di sperimentazione, anche grazie a Cristian Consonni che sul fronte Wikipedia ha sin dall'inizio seguito il lavoro.

La procedura di acquisizione automatica impiegata è chiaramente in fase di perfezionamento, non tutti i problemi sono stati risolti, ma molto è stato fatto e i vantaggi superano di gran lunga gli svantaggi di qualche raro disallineamento semantico. La metodologia richiede comunque alcune verifiche di tipo intellettuale e manuale ed è interessante sottolineare che, in fase di aggiornamento dei dati, se i bibliotecari riscontrano errori possono intervenire, inserendo le correzioni direttamente su Wikidata<sup>38</sup>. Non mancano, come accennato, particolari casistiche su cui si sta lavorando: un *report* della BNCF (al momento a circolazione interna) ne riassume le tipologie e le soluzioni adottabili<sup>39</sup>. Ma è significativo segnalare che, da alcune verifiche a campione condotte su un migliaio di termini, la percentuale di errore dovuta a catture automatiche di link non supera il 10%, un dato per ora molto rassicurante.

Un ulteriore e fondamentale sviluppo nell'interoperabilità fra i due strumenti è stato realizzato a partire dal mese di maggio 2013. Da questa data anche Wikipedia inizia a fornire link inversi al *Nuovo soggettario* implementando un meccanismo di navigabilità reciproca.

**38** Una recente sintesi di quanto sinora realizzato è stata presentata da Maria Grazia Pepe, in occasione dell'assemblea nazionale di Wikimedia Italia, Firenze, 5 aprile 2014: Maria G. Pepe, *Wikipedia Nuovo soggettario. Lo stato dell'arte e gli sviluppi futuri*, <[http://thes.bncf.firenze.sbn.it/Wikipedia\\_NS.pdf](http://thes.bncf.firenze.sbn.it/Wikipedia_NS.pdf)>; <[https://upload.wikimedia.org/wikipedia/commons/o/oe/NuovoSoggettario\\_e\\_Wikipedia\\_05-04-2014.pdf](https://upload.wikimedia.org/wikipedia/commons/o/oe/NuovoSoggettario_e_Wikipedia_05-04-2014.pdf)>.

**39** A scopo esemplificativo, si descrivono alcune possibili casistiche di disallineamento: termini preferiti del *Nuovo soggettario* non accettati in Wikipedia con rinvio ad altri lemmi; termini preferiti del *Nuovo soggettario* linkati a lemmi di Wikipedia che li citano all'interno di lemmi che nel thesaurus non costituiscono forme preferite; termini del *Nuovo soggettario* citati all'interno di lemmi di Wikipedia costituiti da nomi propri non previsti dal thesaurus; termini del *Nuovo soggettario* a cui corrispondono in Wikipedia pagine di disambiguazione; termini del *Nuovo soggettario* in forma non preferita che invece, in Wikipedia, costituiscono lemmi accettati; termini del *Nuovo soggettario* in forma non preferita che in Wikipedia sono citati all'interno di un altro lemma; termini del *Nuovo soggettario* in forma non preferita che anche in Wikipedia sono voci di rinvio ecc.



Figura 3 – Lemma di Wikipedia con link (a fondo pagina) al Nuovo soggettario.

Grazie a questa sincronizzazione, realizzata attraverso il campo P508 di Wikidata, i lettori di Wikipedia, tramite il collegamento esterno disponibile a fondo pagina di ogni lemma, possono transitare non solo nel corrispondente termine del thesaurus ma, grazie al tasto ‘Notizie bibliografiche’ che correda il record del termine, possono anche transitare nell’OPAC della BNCF, visualizzando l’elenco delle opere che la Biblioteca possiede su quel tema. Evidentemente si tratta di un meccanismo che non garantisce le raffinatezze della cosiddetta ‘ricerca a due stadi’, cioè quella che delinea il percorso che conduce dai singoli descrittori alle stringhe di soggetto e da queste alle specifiche notizie bibliografiche a cui sono associate. Ma siamo comunque di fronte ad una novità significativa sul piano culturale, ad un’enciclopedia che, invece di limitarsi ad elencare riferimenti bibliografici a chiusura dei propri lemmi, porta i lettori verso il catalogo di una biblioteca e quindi verso le opere che trattano quell’argomento: come entrare in una biblioteca passando dalle finestre di un’enciclopedia...

Il *Nuovo soggettario* è stato il primo thesaurus multidisciplinare che, tramite il collegamento con DBpedia Italiana, abbia attivato una forma di *interlinking* con una versione di Wikipedia in una specifica lingua, preceduto soltanto da alcune sperimentazioni realizzate, a livello internazionale, in settori specialistici<sup>40</sup>.

La versione italiana di Wikipedia è stata la prima, fra tutte le 287 versioni disponibili, a prevedere collegamenti esterni ‘sistematici’ a un thesaurus interdisciplinare prodotto da una biblioteca. L’annuncio di questa funzionalità è stato dato quasi contemporaneamente dal sito della Biblioteca nazionale centrale di Firenze<sup>41</sup> e da Wikimedia; altre notizie sulla collaborazione sono comparse nella *GLAM Newsletter* di settembre 2013, come Italy report: *Wikimedia Italia meets the National Central Library of Florence*<sup>42</sup>.

Altre biblioteche nazionali, successivamente all’iniziativa della BNCF, si stanno muovendo per collegare i propri strumenti di indicizzazione con le versioni di Wiki-

40 Una delle rare esperienze prototipali di collegamento fra un thesaurus e DBpedia è quella descritta da Joachim Neubert, *Bringing the “Thesaurus for Economics” on to the Web of linked data*, in *Proceedings of the WWW2009 Workshop on Linked Data on the Web, LDOW 2009, Madrid, Spain, April 20, 2009*, <[http://ceur-ws.org/Vol-538/ldow2009\\_paper7.pdf](http://ceur-ws.org/Vol-538/ldow2009_paper7.pdf)>.

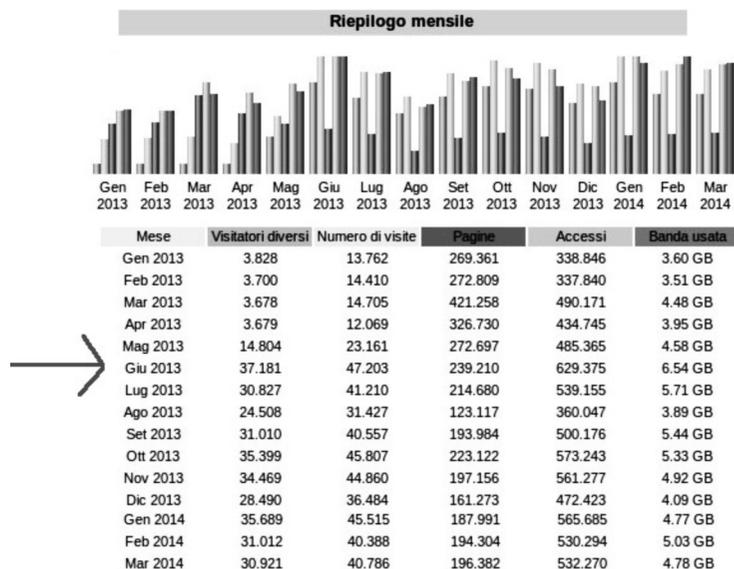
41 Thesaurus Nuovo Soggettario BNCF per Wikipedia, <<http://www.bncf.firenze.sbn.it/notizia.php?id=1195>>.

42 GLAM/Newsletter/September 2013/Contents/Italy report, <[https://outreach.wikimedia.org/wiki/GLAM/Newsletter/September\\_2013/Contents/Italy\\_report](https://outreach.wikimedia.org/wiki/GLAM/Newsletter/September_2013/Contents/Italy_report)>.

pedia nelle rispettive lingue. Ma in nessuno caso, ad oggi, è stata ancora implementata una completa interoperabilità reciproca come nel caso del *Nuovo soggettario*.

La Bibliothèque nationale de France tramite il sistema RAMEAU, la Library of Congress tramite LCSH e la Deutsche Nationalbibliothek con il suo Schlagwort-normdatei, hanno iniziato ad indicare Wikipedia fra le fonti dei propri vocabolari controllati, ma si tratta di citazioni che solo in qualche caso attivano collegamenti, comunque non ancora di tipo reciproco. Ad esempio, dalla voce *Musique de télévision* di RAMEAU si transita alla Wikipedia in versione francese, ma solo alla sua home page<sup>43</sup>. Un *deep link* è attivato a partire dal termine *Caporetto, Bataille de (1917)*, da cui si può arrivare al corrispondente lemma di Wikipedia<sup>44</sup>, ma senza alcuna forma di collegamento inverso.

Il collegamento di Wikipedia con il *Nuovo soggettario* ha comportato un netto incremento delle visite del sito di quest'ultimo. Nel mese di giugno 2013, a pochi giorni dall'attivazione dell'interoperabilità reciproca, queste sono balzate a 37.181, una cifra ben superiore alle 14.804 del mese precedente.



**Figura 4 - Incremento visitatori del *Nuovo soggettario* grazie ai link reciproci con Wikipedia.**

L'incremento ha contagiato anche il fronte di altri servizi forniti dalla Biblioteca ai suoi utenti locali e remoti. Non è un caso, infatti, che nello stesso mese di giugno del 2013 anche le visite del sito della BNCf abbiano registrato una notevole impennata, arrivando a quota 31.245, il 28% in più rispetto a quelle del giugno precedente.

<sup>43</sup> Notice d'autorité RAMEAU, <<http://catalogue.bnf.fr/servelet/autorite?idNoeud=1&ID=11978060&host=catalogue>>.

<sup>44</sup> Notice d'autorité RAMEAU, <<http://catalogue.bnf.fr/servelet/autorite?idNoeud=1&ID=12273273&host=catalogue>>.

Questa esperienza ha dato avvio ad una collaborazione che si è estesa oltre l'ambito del collegamento fra il thesaurus e l'enciclopedia e, soprattutto, ha trovato una formalizzazione ufficiale nella lettera di intenti che il direttore della BNCF e la presidente di Wikimedia Italia hanno sottoscritto il 5 ottobre del 2013. I bibliotecari della Nazionale che si occupano di indicizzazione e di Web semantico e i wikimediani con cui sono in contatto stanno valutando ulteriori sviluppi, indagando la possibilità di raccordare DBpedia con altri *data set* e di consolidare questa collaborazione anche sul piano istituzionale, come sta avvenendo in Francia dove l'avanzamento di DBpédia en français, come si può leggere, «est mené dans le cadre de la plateforme Sémanticpédia dont les partenaires sont: l'équipe Wimmics de Inria, le Ministère de la Culture et de la Communication et l'association Wikimédia France»<sup>45</sup>.

Nel frattempo, a febbraio 2014, i rapporti della BNCF con la DBpedia Italiana e con la Fondazione Bruno Kessler di Trento che la gestisce si sono ulteriormente consolidati. Grazie ai contatti con Cristian Consonni, del *Digital Commons Lab* e con Marco Fossati e Marco Amadori, dell'unità *Web of Data*, è stato concordato l'allineamento dei tempi di aggiornamento delle basi di it.Wikipedia e del *Nuovo soggettario*.



Figura 5 – Logo della DBpedia Italiana e la versione inglese al centro della *Linked Open Data Cloud*.

Un altro sviluppo recente del lavoro italiano riguarda l'idea di usare Wikidata anche su altri fronti come, ad esempio quello del multilinguismo. È in fase di sperimentazione quanto ipotizzato per la prima volta in una discussione svoltasi a Firenze durante la Bibliothakaton del 26 ottobre 2013, dedicata a *Wikipedia, Wikidata e i dati bibliografici*<sup>46</sup>. La BNCF sta lavorando alla possibilità di impiegare Wikidata come *hub* per il recupero automatico di equivalenti inglesi della terminologia messa a disposizione dal *linked data service* di *Library of Congress Subject headings*. Il thesaurus del *Nuovo soggettario* già dal 2006 ha iniziato a corredare alcuni termini dei relativi equivalenti inglesi; i link si attivano tramite il campo 'Equiv. LCSH'. A maggio 2014 i collegamenti ai soggetti in inglese hanno superato quota 5.200, raggiungendo il 25% circa del totale dei termini preferiti del thesaurus. Inoltre, grazie al formato SKOS adottato da entrambi i sistemi di indicizzazione, dal mese di dicembre del 2013, la Library of Congress ha potuto agganciare i collegamenti già attivati dalla BNCF e

<sup>45</sup> DBpédia en français, <<http://fr.DBpedia.org/>>.

<sup>46</sup> Progetto:GLAM/Biblioteche/Bibliothakathon/Firenze, 26 ottobre 2013, <[http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Bibliothakathon/Firenze,\\_26\\_ottobre\\_2013](http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/Bibliothakathon/Firenze,_26_ottobre_2013)>.

inaugurare i link inversi, presentando fra i *Closely Matching Concepts from Other Schemes* anche i descrittori italiani, preceduti dall'icona con relativo flag<sup>47</sup>.

La sfida, come accennato, sarà quella di coinvolgere Wikidata, collegato a tutte le versioni di DBpedia nelle lingue più varie, come piattaforma comune e come *hub* per acquisire collegamenti multilingue fra strumenti di indicizzazione semantica e liste di autorità in lingue diverse, estendendo le funzionalità di Wikidata (integrazione fra Web dei documenti e Web dei dati) al mondo delle biblioteche e dei metadati bibliografici, come ha recentemente illustrato anche Giovanni Bergamin ad un incontro sulle esperienze concrete di alcune biblioteche italiane nell'ambito del Web semantico<sup>48</sup>.

La sperimentazione in corso alla BNCf, infatti, si basa sulla possibilità di acquisire automaticamente link ai soggetti in inglese della *Library of Congress* sfruttando possibili collegamenti tra en.Wikipedia e LCSH e di mapparli con i link che si possono stabilire tra it.Wikipedia e *Nuovo soggettario*. Un progetto che ha potuto beneficiare anche dei risultati ottenuti da un prototipo di mappatura tra LCSH e en.Wikipedia realizzata negli Stati Uniti da Mark Ockerbloom della University of Pennsylvania.

Ulteriori novità riguardano altri sviluppi che la collaborazione fra BNCf e Wikimedia Italia potrebbe favorire. La Biblioteca nazionale di Firenze ha deciso di estendere il proprio coinvolgimento anche a Wikisource, un altro dei progetti della galassia Wikipedia, una sorta di «biblioteca digitale multilingue, che accoglie testi e libri in pubblico dominio o con licenze libere», un progetto suddiviso, come anche Wikipedia, in base a varie versioni linguistiche e che ospita «monografie, documenti, testi di ogni tipologia ed epoca: dalla filosofia alla matematica, dalla letteratura alla numismatica»<sup>49</sup>.

Analogamente a quanto già inaugurato qualche anno fa dalla Bibliothèque nationale de France<sup>50</sup>, anche la BNCf potrà mettere a disposizione di Wikisource alcune

**47** Per una presentazione recente del progetto di interoperabilità fra *Nuovo soggettario* e *Library of Congress Subject Headings*: Anna Lucarelli; Elisabetta Viti, *Florence–Washington round trip: Ways and intersections between semantic indexing tools in different languages*, in *Abstract 27th February 2014, Session 2A. FSR 2014. Faster, smarter and richer. Reshaping the library catalogue. International conference, Roma, 27-28 febbraio 2014*, <<http://www.aib.it/attivita/congressi/fsr-2014/fsr2014-abstract/2014/40480-fsr-session2a-20140227/>>.

**48** Giovanni Bergamin, *SKOS, Nuovo soggettario e Wikidata: Appunti per l'evoluzione dei sistemi di gestione dell'informazione bibliografica*, in *Koha Gruppo italiano, Italian Lectures on Semantic Web and Linked Data: Practical Examples for Libraries, Roma, 7 maggio 2014*, <<https://www.youtube.com/watch?v=fTOfYuwEoho&t=64m30s>>.

**49** Wikisource, <<http://it.wikipedia.org/wiki/Wikisource>>.

**50** Come si può leggere sul sito di fr.Wikisource: «Au travers de ce partenariat, Wikimédia France valorise le travail effectué par les internautes sur le projet Wikisource et démontre que les collaborations entre des projets menés de front par des bénévoles, comme Wikisource et Wikipédia, et des organismes publics comme la BnF, permettent d'ouvrir au plus grand nombre l'accès à des ressources littéraires de qualité. La correction collaborative des textes numérisés hébergés sur Wikisource offre au lecteur une qualité et une fiabilité uniquement possible au travers d'une correction humaine respectueuse du texte original», <[http://fr.Wikisource.org/wiki/Wikisource:Partenariats/Biblioth%C3%A8que\\_nationale\\_de\\_France](http://fr.Wikisource.org/wiki/Wikisource:Partenariats/Biblioth%C3%A8que_nationale_de_France)>. La notizia è riportata anche sul sito di Wikimedia France, <<http://www.wikimedia.fr/wikim%C3%A9dia-france-signe-un-partenariat-avec-la-bnf>>, e fortemente valorizzata, oltre che dai comunicati stampa della BNF, anche dalla stampa nazionale.

digitalizzazioni già realizzate – relative ad opere antiche non gravate dal diritto d'autore – scegliendole secondo criteri da definire (ad esempio, pacchetti tematici di opere a stampa su particolari argomenti), oltreché esaminare la possibilità di includere nella collaborazione anche risorse digitali relative ad alcuni manoscritti.

Dunque, in una logica completamente orientata al *crowdsourcing*, a costi zero e beneficiando della creatività collettiva e collaborativa degli utenti e dei volontari della rete che volessero partecipare, la Biblioteca tramite Wikisource, oltre che diventare partner di un progetto di rilevanza mondiale, consentirà ulteriori possibilità di accesso ad una scelta di risorse digitali già disponibili tramite il proprio sito, valorizzando la fruibilità delle proprie raccolte da parte di utenti remoti e potenziando la consultazione dell'OPAC grazie ai link che si svilupperanno, come già avviene grazie all'interoperabilità fra Wikipedia e il *Nuovo soggettario*.

La partecipazione a Wikisource potrebbe, inoltre, consentire di acquisire miglioramenti e correzioni di eventuali errori che inevitabilmente derivano da trascrizioni automatiche di testi antichi, facilitando la possibilità che le copie digitali diventino il più possibile fedeli alle relative versioni originali. Ma sarà anche un'interessante occasione per sviluppare gli studi sull'indicizzazione automatica di risorse digitali avviata dalla BNCF già qualche anno fa e sinora applicata sperimentalmente solo alle tesi di dottorato di ricerca<sup>51</sup>.

Estendere la collaborazione delle biblioteche (evidentemente non solo delle nazionali) sul fronte di Wikisource, nell'ottica della diffusione di 'beni comuni', potrebbe indurre molte persone a collaborare attivamente per abbattere i costi della digitalizzazione; inoltre, come ha scritto Andrea Zanni il 5 maggio 2014 nella lista citata, progetti di questo tipo consentirebbero, ad esempio, di «pensare ad iniziative in biblioteca, concorsi, collaborazioni con le scuole per far rileggere i libri. Ricordo sempre che con una spesa di 80 euro Wikimedia Italia ha messo in palio tre premi per un concorso di rilettura che hanno portato a più di 4300 pagine rilette e, a quanto dice la Wikimedia Foundation, ad un aumento dei contributori totali. Poi si parla di *low hanging fruit* [...]». Raramente si sono presentate occasioni simili per coinvolgere la società, e in particolare i giovani, nel mondo delle biblioteche e nelle attività che vi si svolgono.

Insomma, possiamo dire che il *flirt* sbocciato fra le biblioteche (anche italiane) e l'universo wikimediano stia davvero diventando un amore corrisposto, ancorché facilitato da rinnovati paradigmi relativi a funzioni, metodologie e obiettivi del lavoro bibliotecario ed anche da forme nuove di comunicazione e di interazione fra contesti professionali, culturali e informativi che, fino a poco tempo fa, sono stati fra loro soltanto amici.

E così questa 'relazione' si sta sviluppando su più piani, alcuni – come abbiamo visto dai progetti della Biblioteca nazionale di Firenze – molto tecnici, basati su ciò che la rete, i nuovi formati di scambio, gli sviluppi delle biblioteche digitali e del Web semantico stanno rendendo possibile. Interessante, a questo riguardo, anche il futuro che riusciranno ad avere funzionalità come quella di Forward to Libraries (FTL), un servizio progettato, anche in questo caso, da Mark Ockerbloom che consente di reperire risorse bibliografiche transitando da una voce di Wikipedia agli OPAC delle

<sup>51</sup> Maria G. Pepe; Elisabetta Viti, *Prime sperimentazioni d'indicizzazione [semi]automatica alla BNCF, in 6° Incontro ISKO Italia, Biblioteca nazionale centrale di Firenze, Firenze, 20 maggio 2013*, <<http://www.iskoi.org/doc/firenze13/pepe&viti.pdf>>.

biblioteche tramite un *template* che crea collegamenti attraverso passaggi da un database (il *tool* è rilasciato con licenza libera) contenente nomi di autori, soggetti, titoli in inglese, derivati da vari cataloghi di biblioteche, e sulla cui fattibilità anche italiana ci si sta appassionatamente confrontando<sup>52</sup>.

In altri casi, i rapporti che si stanno sviluppando sono meno tecnici ma altrettanto interessanti, più legati alla sfera *social*, quella della comunicazione e della condivisione, così che non ci appaiono casuali, né destano ormai meraviglia alcuna, iniziative come quella dello sportello informativo Chiedi al wikipediano (sottile parodia del noto Chiedi al bibliotecario...) che alcuni soci di Wikimedia Italia hanno avuto idea di allestire a Milano il 13 marzo 2014, in occasione del Convegno delle Stelline, per rispondere a domande di bibliotecari su Wikipedia e progetti wiki.

Anche le biblioteche italiane, se lo vorranno, potranno continuare a farsi contaminare dagli aspetti migliori dello spirito collaborativo che, tramite il Web, sta alimentando Wikipedia: potranno acquisirne nuove modalità di scambio delle conoscenze, condividere competenze e, senza rinunciare a svolgere i compiti che sono loro propri, diventare anch'esse una «comunità trasversale», come è stato illustrato da un suggestivo *paper* del già citato convegno milanese<sup>53</sup>.

E così non ci sarà da stupirsi se, in quest'ottica e soprattutto per alcune tipologie di biblioteche che costituiscono componenti significative del sistema culturale del paese, si cominceranno anche a valutare e sperimentare nuove forme di aiuto 'sociale', non certo destinabili a garantirne la sussistenza o i compiti istituzionali, ma magari adatte allo sviluppo di particolari servizi, come una nuova comunità, quella dei sostenitori del *civic crowdfunding*, sta appassionatamente tentando di raccomandarci<sup>54</sup>.

*Un sincero ringraziamento a tutti i colleghi bibliotecari e agli amici wikipediani (non tutti menzionabili nell'articolo) che, in modo aperto e informale, hanno partecipato – con contributi sia teorici che tecnici – alla realizzazione e allo sviluppo dei lavori descritti.*

**52** Il progetto è stato illustrato dal bibliotecario americano il 4 marzo 2013 in *From Wikipedia to our libraries*. Per altre informazioni *Forward to libraries*, <[http://wiki.wikimedia.it/wiki/Forward\\_to\\_libraries](http://wiki.wikimedia.it/wiki/Forward_to_libraries)>. Altro suo contributo (successivo di qualche mese) sul tema: John Mark Ockerbloom, *Forward to libraries: Experiences connecting digital libraries, local libraries, and Wikipedia*, in *Digital Library Federation Forum, Austin, novembre 2013*, <[http://works.bepress.com/john\\_mark\\_ockerbloom/16](http://works.bepress.com/john_mark_ockerbloom/16)>.

**53** Pierfranco Minsenti, *Wikipedia come ambiente di lavoro collaborativo per una comunità trasversale alle biblioteche*, in *La biblioteca connessa. Come cambiano le strategie di servizio al tempo dei social network*, Convegno delle Stelline, Milano, 13-14 marzo 2014, abstract: <<http://www.convegnostelline.it/relatori1.php?IdUnivoco=69>>.

**54** Su questo tema, qualche giorno fa, un ricercatore del Center for Civic Media, istituto che opera presso il Massachusetts Institute of Technology (MIT), Rodrigo Davies, *Civic Crowdfunding. Four Things We Know, Two Things We Don't*, 9 maggio 2014, <[http://rodrigodavies.com/blog/2014/05/09/civic-crowdfunding-four-things-we-know-two-things-we-dont.html?utm\\_content=bufferb750e&utm\\_medium=social&utm\\_source=twitter.com&utm\\_campaign=buffer](http://rodrigodavies.com/blog/2014/05/09/civic-crowdfunding-four-things-we-know-two-things-we-dont.html?utm_content=bufferb750e&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)>.

Articolo proposto il 23 maggio 2014 e accettato il 27 giugno 2014.

---

**ABSTRACT** AIB studi, vol. 54 n. 2/3 (maggio/dicembre 2014), p. 241-259. DOI 10.2426/aibstudi-10108.

---

ANNA LUCARELLI, Biblioteca nazionale centrale Firenze, piazza Cavalleggeri 1/A, 50122 Firenze, e-mail [anna.lucarelli@beniculturali.it](mailto:anna.lucarelli@beniculturali.it).

**«Wikipedia loves libraries»: in Italia è un amore corrisposto...**

Nell'ambito del Progetto GLAM, la comunità di Wikipedia ha sviluppato la campagna *Wikipedia Loves Libraries* con l'obiettivo di mettere in comunicazione l'enciclopedia – ricca di milioni di risorse informative – con il mondo delle biblioteche, degli archivi e dell'open access. Sulla scia di quanto sta avvenendo in altri Paesi, anche in Italia si stanno diffondendo iniziative che hanno coinvolto le biblioteche e i loro utenti. Le prime attività sono iniziate circa due anni fa grazie ad alcuni bibliotecari ed appassionati di Wikimedia Italia. In un lasso di tempo brevissimo sono stati realizzati eventi e organizzati seminari, conferenze e corsi di formazione; sono state siglate convenzioni ed elaborati progetti che hanno coinvolto anche l'ambito della catalogazione e del controllo d'autorità. L'ICCU ha avviato un lavoro per valorizzare la fruizione dei record di autorità prodotti nell'ambito di SBN, intensificando la partecipazione italiana al VIAF (Virtual International Authority File) e il collegamento alle voci di Wikipedia. La Biblioteca nazionale centrale di Firenze ha attivato link fra i termini del Thesaurus del Nuovo Soggettario e i corrispettivi lemmi di Wikipedia, e a partire dal maggio 2013 Wikimedia Italia ha favorito l'implementazione dei link reciproci (interlinking). In questo modo è stato realizzato il primo prototipo di interoperabilità fra Wikipedia e uno strumento di indicizzazione per soggetto multidisciplinare prodotto nel mondo delle biblioteche. La BNCF sta lavorando anche per sperimentare l'uso di Wikidata per potenziare i collegamenti con i soggetti in inglese della Library of Congress, e per estendere la collaborazione al progetto Wikisource.

Tutto insomma fa pensare che le biblioteche italiane possano iniziare ad acquisire lo stesso spirito aperto, creativo e collaborativo che, tramite il web, sta alimentando Wikipedia.

**«Wikipedia loves libraries»: a reciprocal love in Italy...**

As part of the GLAM projects, the Wikipedia community has launched the *Wikipedia Loves Libraries* campaign aiming to connect Wikipedia with the world of libraries, archives and OA. In the last couple of years events were carried out, workshops, conferences and training courses were organized, agreements were signed and projects – which also involve the cataloguing and authority control field – were developed. ICCU (Central Institute for Union Catalogue) has enhanced the use of the National Library Service (SBN) authority records by increasing the Italian involvement in the VIAF (Virtual International Authority File) and its connection to Wikipedia entries. The National Central Library of Florence has activated links between the Nuovo Soggettario Thesaurus and Wikipedia, and since May 2013 Wikimedia Italia promoted the implementation of interlinking, thus carrying out the first prototype of interoperability between Wikipedia and a multidisciplinary subject indexing tool produced by libraries. The BNCF is also testing the use of Wikidata to improve the links to the Library of Congress subject headings, and to extend the collaboration to the Project Wikisource.

Everything suggests that Italian libraries can begin to acquire the same open, creative and collaborative attitude that's sustaining Wikipedia.