

Nuove scommesse della BNCF: Wikipediani in residence, Wikisource e altro ancora

Anna Lucrelli

Biblioteca nazionale centrale Firenze

Il 28 novembre 2014 la Biblioteca nazionale di Firenze (BNCF) ha ospitato il convegno “Promuovere la cultura. Sfide e alleanze tra Biblioteche e Wikipedia”. L’iniziativa si è inserita nel percorso di collaborazione che la Biblioteca ha avviato da oltre un anno con la Fondazione Wikimedia Italia e ha rappresentato un momento importante e di larga partecipazione. Si è confermata la progressiva conoscenza, ormai anche in Italia, di GLAM-Wiki Initiative, il progetto che intende favorire la condivisione delle risorse di istituzioni depositarie delle memorie culturali di vari Paesi del mondo (Galleries, Libraries, Archives, and Museums), con l’universo wikipediano¹. Non a caso, il convegno è stato inaugurato dall’australiano Liam Wyatt, già wikipediano *in residence* al British Museum ed oggi coordinatore di European GLAM-Wiki. L’incontro, il secondo su questi temi ospitato dalla BNCF, dopo il *Bibliothackaton* dell’ottobre 2013, è stato un’occasione preziosa per condividere i risultati delle esperienze italiane realizzate negli ultimi mesi in ambiti bibliotecari, museali e documentari². Quanto segue costituisce l’intervento di chi scrive al suddetto convegno.

Come dimostrano alcune esperienze straniere già realizzate, anche le biblioteche nazionali possono giocare un ruolo fondamentale nello

scenario GLAM. Sarebbe sbagliato pensare che esse, per la loro storia, natura e compiti istituzionali, debbano rimanerne escluse. Nonostante siano tradizionalmente meno esposte alla dimensione pubblica e partecipativa tipica del Wiki, anche le biblioteche nazionali sono coinvolte nella “trasversalità” verso cui il web sta sempre più orientando l’universo



Anna Lucrelli al Convegno “Promuovere la cultura. Sfide e alleanze tra Biblioteche e Wikipedia” (BNCF - 28 novembre 2014)

¹ <http://en.wikipedia.org/wiki/Wikipedia:GLAM>.

² http://it.wikipedia.org/wiki/Progetto:GLAM/Biblioteche/progetti/Sfide_e_alleanze_tra_Biblioteche_e_Wikipedia.

della conoscenza e i suoi archivi. Proprio per questo potranno assumere un ruolo di leadership grazie ai linked data che danno un valore più produttivo e aperto ai contenuti e ai dati che esse producono. Se sapranno convertire e rendere disponibili nei linguaggi del web semantico i loro dati, rendendoli interoperabili, le Nazionali potranno dare un senso nuovo alle attività che svolgono, diffondendo il risultato del proprio lavoro in contesti di open access, al di fuori del veicolo esclusivo e tradizionale dei cataloghi e servendo un numero di cittadini ben superiore a quello attuale. Questo scenario offre alle biblioteche nazionali un'opportunità unica per dimostrare il loro valore pubblico, per giustificare gli investimenti impiegati nella digitalizzazione delle loro risorse e, come ha ricordato più volte Gildas Illien (direttore del Département de l'Information bibliographique et numérique, presso la Bibliothèque nationale de France), per restituire benefici alla società e ad altri *data set* della pubblica amministrazione³. Per realizzare tutto questo la collaborazione ai vari progetti di Wikimedia sarà preziosa, una collaborazione peraltro favorita da una concezione rinnovata degli obiettivi del lavoro bibliotecario ed anche da forme nuove di interazione fra contesti professionali, culturali e informativi che finora erano amici, ma solo a distanza. Un processo che sta contagiando anche l'Italia, dove l'attenzione e l'interesse delle biblioteche per l'universo wikipediano si sta diffondendo con una certa intensità⁴.

In questo percorso la BNCF sta avendo un ruolo attivo. Il suo coinvolgimento è iniziato circa due anni fa grazie al progetto del *Nuovo soggettario* e da lì si è diffuso ad altri ambiti, attraverso tappe serrate e straordinariamente veloci.

Nel gennaio 2013, chi scrive è entrata a far parte di una lista di discussione appositamente allestita da wikipediani per dialogare con bibliotecari⁵, manifestando l'interesse della BNCF a rendere più disponibili in forma aperta i metadati catalografici prodotti dalla Biblioteca e ad approfondire l'implementazione dei collegamenti fra la versione italiana di *Wikipedia* e il Thesaurus del *Nuovo soggettario* (già disponibile, dal 2010, in formati adeguati al web semantico e con metadati esposti in triple RDF)⁶. Dopo solo due mesi, una procedura predisposta dal Dipartimento Servizi informatici della Biblioteca, ha consentito al Thesaurus del *Nuovo soggettario* di acquisire automaticamente link ai corrispondenti lemmi di *Wikipedia*, link che sono passati così dai 990 inseriti manualmente fino a quel momento ad oltre 12.000. Due mesi ancora (maggio 2013) e i collegamenti diventano reciproci: in fondo a molti lemmi della *Wikipedia* in italiano, i lettori trovano link al Thesaurus attraverso i quali possono transitare nell'Opac della BNCF, arrivando alle specifiche notizie bibliografiche sull'argomento⁷. Il *Nuovo soggettario* è stato il primo thesaurus multidisciplinare che, sfruttando il collegamento con DBpedia (la base di dati strutturati estratti da *Wikipedia*)⁸, ha attivato

³ Gildas Illien, *Interview with Gildas Illien*, «Information Standards Quarterly», 25 (2013), n. 4, p. 22-29. <http://www.niso.org/apps/group_public/download.php/11938/IP_Illien_BnF_isqv25no4.pdf>.

⁴ Anna Lucarelli, *Wikipedia loves libraries: in Italia è un amore corrisposto...* «Aib studi», 54 (2014), n. 2/3, p. 241-259.

⁵ bibliotecari@wikimedia.it.

⁶ <http://thes.bncf.firenze.sbn.it/thes-dati.htm>.

⁷ Al lavoro di *matching* svolto in BNCF ha dato un contributo anche Cristian Consonni, di Wikimedia Italia. Una descrizione di questo lavoro prototipale è stata fornita da Maria Grazia Pepe, *Wikipedia e Nuovo soggettario: lo stato dell'arte e gli sviluppi futuri. Assemblea nazionale di Wikimedia Italia*. Firenze 5 aprile 2014. <http://upload.wikimedia.org/wikipedia/commons/0/0e/NuovoSoggettario_e_Wikipedia_05-04-2014.pdf>.

⁸ <http://it.dbpedia.org/>.

una forma di *interlinking* con una versione di Wikipedia in una specifica lingua. Anche grazie ai lavori prototipali descritti, a ottobre del 2013 la BNCF e Wikimedia Italia hanno firmato una convenzione incentrata da un lato sugli sviluppi dell'interoperabilità fra il *Nuovo soggettoario* e *Wikipedia*, dall'altro sulla possibile condivisione di altri *data set* gestiti da entrambe le istituzioni. Il lavoro si è via via esteso su vari fronti e, tra l'altro, ha prodotto un'importante sperimentazione sulla possibilità per il Thesaurus di acquisire automaticamente equivalenti inglesi previsti da *Library of Congress subject headings* (LCSH), tramite le funzionalità multilingue di Wikidata⁹.

La convenzione fra BNCF e Wikimedia Italia si è ulteriormente perfezionata a giugno del 2014, quando una lettera di intenti ha esteso la collaborazione anche a Wikisource, la biblioteca digitale multilingue, nata oltre dieci anni fa con lo scopo di creare un archivio di testi di tutte le discipline non coperti da diritto d'autore, di pubblico dominio o con licenze libere. Tali testi, è questa l'assoluta novità, sono verificabili e correggibili (ad esempio nel caso di errori dovuti all'OCR) da una comunità

di lettori volontari. A novembre del 2014, una prima risorsa digitale della BNCF diventa disponibile tramite Wikisource.

Wikisource è un progetto strategico della galleria wikipediana al quale istituzioni nazionali di tutto il mondo, sia bibliotecarie che archivistiche, partecipano già da tempo¹⁰. Celebri i casi dei National Archives, agenzia del governo statunitense che ha lo scopo di facilitare l'accesso ai documenti depositati negli archivi delle istituzioni federali¹¹, così come della Gerald Ford Presidential Library¹² o della Bibliothèque Nationale de France (BNF), che ha avviato la collaborazione attraverso 1.400 digitalizzazioni già disponibili tramite Gallica, ottenendo risonanza e grande interesse nel mondo della cultura e dell'informazione francese¹³.

Per la Biblioteca nazionale di Firenze la partecipazione a Wikisource è nata con l'obiettivo di ampliare, a costo zero, la visibilità e la fruibilità delle proprie risorse digitali, mettendo a disposizione opere non gravate dal diritto d'autore. I vantaggi che ne conseguono sono di varia natura: diventare partner di un progetto di rilevanza mondiale; utilizzare forme

⁹ http://www.wikidata.org/wiki/Wikidata:Main_Page. Cfr., su questi temi, le relazioni al convegno fiorentino del 28 novembre 2014: Luca Martinelli, *Wikidata: un nuovo modo di diffondere dati strutturati*; Giovanni Bergamin, *La sfida di Wikidata alle biblioteche*; Elisabetta Viti, *Wikidata: nuove opportunità di multilinguismo nel Thesaurus del Nuovo soggettoario*.

¹⁰ http://it.wikisource.org/wiki/Pagina_principale.

¹¹ <http://www.archives.gov/>; sulla partecipazione a Wikisource, vedi: https://en.wikisource.org/wiki/Wikisource:WikiProject_NARA.

¹² https://en.wikisource.org/wiki/Wikisource:WikiProject_Gerald_R._Ford_Presidential_Library.

¹³ "Wikimédia France et la Bibliothèque nationale de France ont signé un accord de partenariat qui va permettre à tous les internautes, au travers de Wikisource, d'avoir accès aux transcriptions d'œuvres tombées dans le domaine public issues de Gallica. Au final, ce sont 1400 textes en français qui vont être intégrés à Wikisource. Ces textes, numérisés par la BnF, ont bénéficié d'une transcription automatique afin de permettre la recherche directe dans le texte. Toutefois, sur des textes anciens, cette reconnaissance automatique introduit souvent des erreurs. Ce partenariat va permettre aux internautes de participer à la correction des textes afin que ceux-ci soient en tout point conformes avec la version originale. 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", https://fr.wikisource.org/wiki/Wikisource:Partenariats/Biblioth%C3%A8que_nationale_de_France. Per l'elenco dei progetti di partenariato fra Wikisource e istituzioni di tutto il mondo: <https://fr.wikisource.org/wiki/Wikisource:Partenariats>.

di *crowdsourcing*, avvalendosi della “creatività collettiva” per la correzione dei testi digitalizzati; incrementare la fruibilità delle raccolte da parte di utenti remoti, con una maggiore consultazione del proprio Opac grazie ai link sviluppati, come già avviene tramite l’interoperabilità fra *Wikipedia* e *Nuovo soggettario*. Vantaggi ai quali potranno aggiungersene altri di natura più prettamente catalografica, ad esempio la possibilità di sviluppare sperimentazioni sul trattamento di risorse digitali come quella già avviata sull’impiego del Thesaurus per progetti di indicizzazione automatica e semi-automatica.

La scelta di partecipare a Wikisource ha stimolato la BNCf ad analizzare le varie tipologie di documenti (manoscritti, materiali grafici, opere a stampa) già archiviati nella propria teca digitale, dandole l’occasione per riflettere sia sui criteri di scelta usati, sia sugli aspetti tecnici inerenti a formati e modalità di conservazione. La riflessione naturalmente ha tenuto conto di quanto realizzato dalla Biblioteca con la partecipazione a noti progetti come Google Books, Proquest¹⁴ e Magazzini digitali, progetto quest’ultimo che ha implementato “un sistema per la conservazione permanente dei documenti elettronici pubblicati in Italia e diffusi tramite rete informatica, in attuazione della normativa sul deposito legale”¹⁵. Naturalmente era importante chiedersi quali opere e edizioni fosse meglio scegliere per Wikisource, se privilegiare quelle non ancora disponibili sul web in altri archivi, oppure quelle particolarmente rappresentative della nostra cultura nazionale; opere di larga diffusione o, al contrario, di interesse specialistico;

risorse in particolari formati piuttosto che in altri. Tutte tematiche che, come è noto, ne coinvolgono altre ben più ampie e legate alle *policies* delle biblioteche digitali, un fronte sul quale, malgrado i progetti già avviati da tempo (vedi ad esempio TELplus)¹⁶, le biblioteche continuano a comunicare troppo poco. Le conseguenze sono i frequenti fenomeni di ridondanza operativa, la sovrapposizione di progetti e, non ultima, la duplicazione di spese e dunque gli sprechi di risorse finanziarie. Nel caso specifico della partecipazione della BNCf a Wikisource, non è stato semplice individuare edizioni non gravate da diritti e non già disponibili, ad esempio, tramite Google Books, Proquest, o la biblioteca digitale di Liberliber¹⁷; è ovvio che, più un’opera è considerata importante, maggiore è la probabilità che ne esistano già una o più digitalizzazioni. Dal punto di vista tecnico, una problematica altrettanto importante ha riguardato il tipo di immagini da fornire rispetto alle modalità di digitalizzazione e archiviazione impiegate dalla Biblioteca. Nel caso in cui non fosse stato pronto, sarebbe stato necessario rielaborare il master digitale del testo scelto. Non si doveva dimenticare, inoltre, che obiettivo di Wikisource è creare OCR da indicizzare, mentre scopo della BNCf è di offrire la consultazione sul web delle proprie digitalizzazioni e contemporaneamente l’archiviazione di master digitali ad una risoluzione molto alta, finalizzata essenzialmente alla conservazione e tale da soddisfare qualunque tipo di rielaborazioni della copia digitale come, ad esempio, quelle di tipo editoriale¹⁸. Altro aspetto da non sottovalutare, infine, era lo spirito del proget-

¹⁴ <http://www.bncf.firenze.sbn.it/notizia.php?id=1268>. Per una sintesi su questi progetti della BNCf, vedi Pierantonio Metelli, *La digitalizzazione di massa: Google e Proquest*. In *1861/2011: l’Italia unita e la sua Biblioteca. [Catalogo della Mostra tenuta a Firenze]*, Firenze: Polistampa, 2011, p. 205-206.

¹⁵ <http://www.bncf.firenze.sbn.it/pagina.php?id=212&rigamenu=Magazzini%20Digitali>.

¹⁶ <http://www.theeuropeanlibrary.org/tel4/aboutus>.

¹⁷ <http://www.liberliber.it/>.

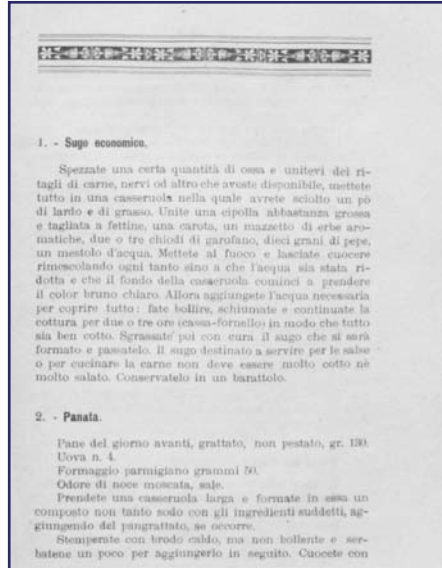
¹⁸ *Technical guidelines for digital cultural content creation programmes*, Version 1.0, edited by Pete Johnston, David Dawson, Roma: MINERVA, 2004-2005. <www.minervaeurope.org/structure/workinggroups/servprov/documents/technicalguidelines1_0.pdf>.

to Wikisource, che non è solo quello di facilitare l'accesso a una risorsa digitale ma anche e soprattutto, come già accennato, di coinvolgere la sua comunità nella ri-lettura del testo. E così, sulla base di questo ventaglio di considerazioni, la scelta della BNCF è stata improntata a criteri diversi, ad esempio, da quelli impiegati dalla BNF, la quale per Wikisource ha privilegiato "opere simbolo" della cultura nazionale francese come l'*Encyclopedie* di Diderot e D'Alembert nella famosa edizione del 1751¹⁹.

La nostra decisione, invece, è stata quella di scegliere un testo semplice, fruibile da tutti, di sole 73 pagine, su una sola colonna, con poche illustrazioni ma di buona qualità digitale. Si è voluto onorare l'anniversario della Prima guerra mondiale e mettere a disposizione un'opera già fruibile su Internet ma solo tramite Europea la quale, come è noto, contiene materiali di natura molto diversa fra loro, comprese risorse disponibili anche solo in formato immagine. Così la collaborazione della BNCF a Wikisource è iniziata con un'edizione del

1916, caratterizzata da una grafica con accenti liberty²¹.

Si tratta di un manuale di ricette per il tempo di guerra, da preparare con i pochi mezzi di cui si disponeva. Minestre con le chiare d'uo-



Ricette per il tempo di Guerra



Risorsa digitale della BNCF disponibile anche su Wikisource

¹⁹ http://fr.wikisource.org/wiki/Encyclop%C3%A9die,_ou_Dictionnaire_raisonn%C3%A9_des_sciences,_des_arts_et_des_m%C3%A9tiers.

²⁰ Sulla partecipazione della BNCF a Europea: <<http://www.bncf.firenze.sbn.it/pagina.php?id=217&rigamenu=Europeana>>.

²¹ https://it.wikisource.org/wiki/Indice:Manuale_150_ricette_di_cucina_di_guerra.djvu.

vo, sugli economici, pappe per bimbi convalescenti, ecc.

Come si legge nell'introduzione, un libro pubblicato "per venire in aiuto alle donne italiane costrette dalle esigenze della guerra a modificare le abitudini delle loro famiglie" e, come indicato in quarta di copertina, venduto "a totale beneficio del Comitato nazionale di assistenza ai mutilati di guerra".

I primi vantaggi della condivisione di questo testo su Wikisource non si sono fatti attendere. In tempi brevissimi, a soli due giorni dalla sua disponibilità, è iniziata la ri-lettura collaborativa, cosa che peraltro ha consentito di scoprire che mancavano due pagine dell'indice, permettendo di ri-digitalizzarle e integrarle anche nella teca BNCf.

Per la Biblioteca si tratta, naturalmente, di un inizio, di una prima sperimentazione cui farà seguito la messa a disposizione di altre opere e di altre tipologie di risorse, oltre che l'allargamento del progetto ad alcuni manoscritti di cui sia disponibile la trascrizione. Le riflessioni sui criteri di scelta continueranno ad essere

approfondite, anche nella consapevolezza che nella problematica della selezione dei testi da privilegiare sono coinvolte questioni di interesse filologico e testuale²².

La decisione della Biblioteca nazionale di Firenze di estendere la collaborazione già avviata sul fronte di Wikipedia anche al progetto Wikisource, potrà dunque favorire nuove e innumerevoli possibilità, da cui trarrebbero benefici, non soltanto il sistema dei servizi bibliografici italiani, ma un nuovo e più allargato panorama di fruitori.

Il coinvolgimento della Biblioteca nei progetti di Wikimedia Foundation è destinato ad allargarsi, probabilmente anche prevedendo la possibilità di attivare collaborazioni con i cosiddetti wikipediani *in residence* (WIR). Come è noto, istituzioni culturali e bibliotecarie di tutto il mondo, in alcuni casi molto prestigiose per la loro storia, dal 2010 accolgono giovani che possano facilitare i loro contatti e la cooperazione con la comunità wikipediana e con i progetti di Wikimedia. I WIR promuovono incontri e gruppi di lavoro, studiano come otti-



La pagina su Wikipedia (in inglese) dedicata ai Wikipediani in residence

²² Pierfranco Minsenti, Università IUAV di Venezia, ha sollevato, durante il convegno, la questione di come Wikisource debba affrontare i criteri in base ai quali preferire un'edizione di un'opera rispetto ad un'altra: si tratta di un tema che riguarda in realtà i progetti di digitalizzazione in generale, spesso purtroppo avulsivi da problematiche di critica testuale e filologia dei testi a stampa. A convegno terminato, il confronto di opinioni si è prolungato nella lista di discussione citata. Vedi anche, su questa stessa rivista, Federico Meschini, *Edizioni critiche digitali: sul rapporto tra testo, edizione e tecnologia*, «DigItalia», vol. 2 (2013), p. 24-42, <<http://digitalia.sbn.it/article/view/829/554>>.

mizzare la fruibilità dei dati di una biblioteca e come renderli disponibili in formati adeguati al semantic web. Il loro contributo è stato fondamentale in certi contesti: non è un caso che OCLC, proprio grazie ad un wikipediano *in residence*, Max Klein, abbia avviato la condivisione di dati fra la versione inglese di Wikipedia e il progetto VIAF (Virtual International Authority File). L'elenco di coloro che hanno svolto presso istituzioni prestigiose l'attività di wikipediani *in residence* è ormai molto nutrito: da Liam Wyatt, primo a svolgere questa attività presso il British Museum, ad innumerevoli altri presso biblioteche nazionali di tutto il mondo, fino alle recenti esperienze italiane presso la Biblioteca civica di Mestre, il Mart di Rovereto e la BEIC²³.

Se accolto in BNCF, un wikipediano *in residence* potrebbe collaborare ad adempiere agli obblighi degli Open data, a sperimentare la conversione dei record bibliografici in formato RDF (migliorando la presenza della Biblioteca sul web), a valorizzare l'enorme quantità di metadati che produce rendendoli più fruibili e, naturalmente, ad ottimizzare sempre più i collegamenti fra *Nuovo soggettario* e *Wikipedia*. Non ci sono dubbi che il coinvolgimento delle

biblioteche con l'universo wikipediano si stia basando, a livello tecnico, su quanto i nuovi formati di scambio e gli sviluppi del web semantico stanno rendendo possibile, ma contemporaneamente anche sullo spirito collaborativo e *social* dei progetti wiki.

Sarà molto utile iniziare a mescolarsi davvero con professionalità e competenze diverse, nella consapevolezza che gli sviluppi tecnologici degli ultimi anni stanno finalmente favorendo la convergenza fra le identità professionali delle biblioteche, dei musei e degli archivi. Sono d'accordo con chi sostiene, ad esempio, che l'approccio "sintetizzante" e "contestualizzante" di cui sono capaci gli archivisti e i curatori di musei possa essere utile ai bibliotecari per ripensare il senso delle descrizioni bibliografiche, degli authority records (troppo poveri per le esigenze dei nuovi utenti?), per interrogarsi su come integrare e arricchire dati, o creare accessi di tipo tematico alle digitalizzazioni, come stanno ormai realizzando i nuovi portali della British Library e della BNF²⁴. Risultati che le biblioteche nazionali potranno raggiungere in maniera efficace proprio confrontandosi con l'universo dei linked open data e con la galassia di *Wikipedia*.

²³ http://outreach.wikimedia.org/wiki/Wikipedian_in_Residence#List_of_Wikipedians_in_Residence.

²⁴ Tematiche proposte alcuni mesi fa da Pierfranco Minsenti nella Lista bibliotecari@wikimedia.it.

L'ultima consultazione dei siti Web è avvenuta nel mese di dicembre 2014.