

Qualche domanda sulla digitalizzazione

Conversando con Valeria De Francesca e Federica Viazzi

LUCIA SARDO

Alma Mater Studiorum Università di Bologna
Campus di Ravenna
lucia.sardo@unibo.it

La collana Library Toolbox della Editrice Bibliografica si arricchisce di un nuovo contributo¹ dedicato alla gestione delle collezioni digitalizzate.

In questo caso specifico le autrici, forti dell'esperienza maturata in anni di lavoro presso la BEIC, presentano al lettore, con encomiabile capacità di sintesi, le principali attività indispensabili per la gestione di una collezione di libri digitalizzati.

Poiché la collana impone limiti di spazio, ed è giusto che sia così, l'argomento viene trattato senza tralasciare le indicazioni principali; tuttavia non potendo entrare nel dettaglio, alcune curiosità vengono lasciate senza risposta.

È per questo che, invece di una semplice recensione, si è preferito cercare di approfondire e di dare una risposta alle curiosità emerse durante la lettura, grazie a un dialogo con le autrici, con l'augurio che siano un interessante completamento del volume e che possano fornire ulteriori spunti di riflessione per coloro che a vario titolo si devono impegnare o sono impegnati in progetti di questo tipo.

La digitalizzazione di libri è una attività ormai entrata a pieno titolo fra quelle "normalmente" svolte dalle biblioteche, con diversi livelli di coscienza, attenzione e con alterni risultati finali.

Tutto ciò si traduce, inevitabilmente, in grandi portali o in piccole raccolte frammentarie, in progetti in collaborazione con privati e aziende o in attività "fai

da te", in progetti strutturati che hanno solide basi progettuali o in estemporanei tentativi di rendere più visibile il posseduto della biblioteca.

La presenza di linee guida, di normative, di documenti di buone pratiche aiuta sicuramente le biblioteche che avviano progetti di digitalizzazione, ma questo spesso non basta per poter poi arrivare a un buon risultato finale, che dovrebbe essere la facilità di accesso alla risorsa digitalizzata e la sua altrettanto facile individuazione, grazie innanzitutto a un corretto processo di catalogazione e metadateazione.

I grandi progetti che coinvolgono realtà istituzionali importanti non sono quelli di cui vorremmo occuparci in questa breve conversazione, piuttosto l'attenzione vorrebbe focalizzarsi (grazie all'esperienza delle autrici nella realizzazione di una biblioteca esclusivamente digitale) su quelle che possono essere le necessità di realtà di media dimensione, con importanti raccolte ma non sempre in grado di avere tutte le competenze, le risorse, il tempo necessari per svolgere al meglio queste attività.

Le questioni su cui ho chiesto alle autrici di focalizzare l'attenzione riguardano in primo luogo le fasi preliminari di un progetto di digitalizzazione, ovvero cosa serve per avviare un progetto, quindi le modalità di visualizzazione dei volumi digitalizzati, che a mio avviso non è secondaria o marginale e infine la questione della conservazione, centrale, in quanto non

riguarda solo l'oggetto digitalizzato in sé, ma anche i metadati che lo accompagnano e lo spazio dove viene conservato.

Il tutto con un approccio molto pratico, con la consapevolezza che quanto viene detto può diventare obsoleto in tempi relativamente brevi.

La prima domanda, apparentemente banale ma decisiva, è quella che riguarda il cosa fare per creare una collezione di libri digitalizzati. Quali scelte le biblioteche devono affrontare per selezionare i materiali da digitalizzare, considerando che le raccolte sono ricche di risorse potenzialmente interessanti?

Il problema non è trovare cosa digitalizzare, ma scegliere oculatamente ciò che si vuole rendere disponibile anche online.

Il processo richiede una pianificazione attenta che analizzi tutti gli aspetti che vedranno coinvolta la raccolta: dalle fasi preliminari fino alla valorizzazione una volta conclusi i lavori, senza ovviamente trascurare il problema dell'accessibilità sul lungo periodo.

Il motivo per cui è fondamentale programmare con attenzione le attività prima di avviare il progetto è abbastanza ovvia: innanzitutto i costi, in termini di risorse umane ed economiche, tutt'altro che irrilevanti, rendono importante non sprecare nemmeno un centesimo di quanto messo a budget; spesso poi avere una visione sul medio-lungo periodo diventa fondamentale per la riuscita del progetto, considerando che molti progetti "accessori" a quella che è l'ordinaria amministrazione delle biblioteche vengono finanziati attraverso bandi o contributi straordinari che esigono una rendicontazione dettagliata.

Le domande che si devono porre i bibliotecari, o comunque i professionisti che intendono approcciare una collezione digitale, e alle quali è necessario trovare una risposta prima di avviare un progetto di digitalizzazione, si possono ricondurre alla regola del giornalismo anglosassone delle 5 W, o meglio ancora agli "otto elementi" di San Tommaso d'Aquino, attinti dal linguaggio filosofico e approdati nella nostra quotidianità in questa epoca di *post-truth*.

Su cosa dobbiamo ragionare, quindi, a tavolino, prima di cominciare a creare la nostra collezione?

Sembra banale, ma la selezione dei materiali va effettuata valutandone lo stato di conservazione. Per esempio: se vogliamo avviare una digitalizzazione che preservi i testi più preziosi dalle consultazioni, dobbiamo valutare se si ha a disposizione uno staff con le

competenze necessarie a maneggiare documenti così delicati.

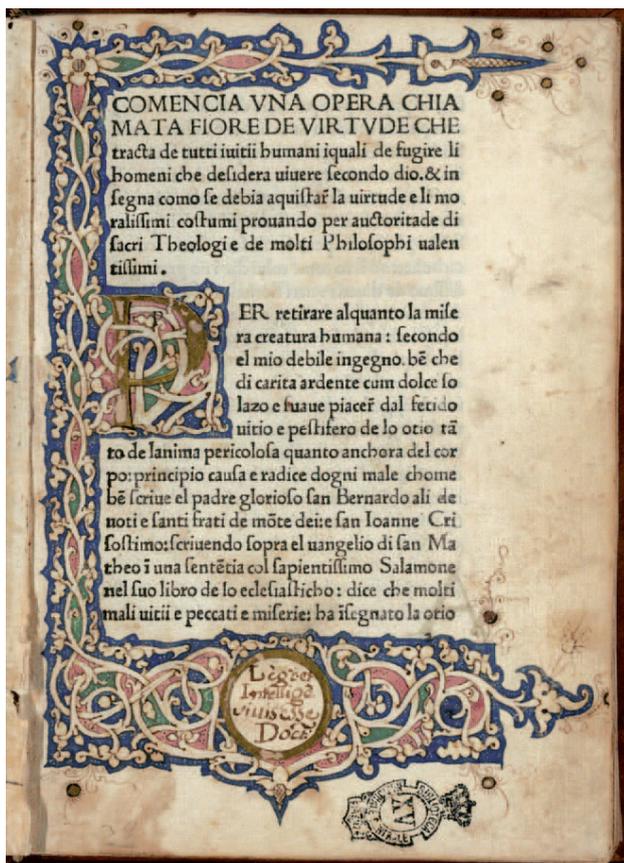
Altro elemento da tenere in considerazione per la scelta dei materiali è il pubblico al quale ci si vuole rivolgere: scegliere di digitalizzare opere che possano interessare, per esempio, un'utenza specialistica, richiederà un flusso di lavoro diverso rispetto a quello necessario per mettere a disposizione le annate del periodico locale a partire dalla sua fondazione. Stesso discorso vale per la valorizzazione e la disseminazione delle risorse successive alla creazione della collezione: pensare di avviare *partnership* di ampio respiro, quale, per esempio, un progetto GLAM in collaborazione con la Wikimedia Foundation, apre prospettive diverse rispetto a sapere di poter pubblicizzare i risultati ottenuti solo sul portale dell'ente di afferenza.

Dopo aver trovato "cosa" vogliamo digitalizzare, dobbiamo chiederci "come".

La selezione dei materiali è legata a doppio filo, in un rapporto dialettico, con gli strumenti e le risorse che abbiamo a disposizione. Occorre infatti essere consapevoli di quali sono gli strumenti su cui possiamo contare: scanner per le prime fasi, computer e relativa potenza di calcolo, spazio di archiviazione; bisogna chiedersi: quante e quali risorse umane si potranno dedicare al progetto e, non ultimo, di quanta disponibilità economica beneficerà il progetto? Consisterà in un'elargizione *una tantum* o invece avrà un sostegno garantito? Il "come" ci porta direttamente al "chi": nel recente Convegno delle Stelline abbiamo sentito parlare di *digital mind transformation*, possiamo semplificare all'estremo il concetto (che richiederebbe un saggio indipendente) con la frase "perché fare da soli ciò che si può fare insieme?". Nel considerare tutte le variabili e definire un progetto di digitalizzazione delle nostre collezioni è importante considerare la possibilità di attivare delle collaborazioni, così da poter rendere scalabili i costi e condividere le risorse e le idee.

Un'ulteriore domanda, forse la più difficile alla quale fornire una risposta, è "perché". Qual è il valore aggiunto che la nostra collezione digitale porta all'ecosistema culturale?

A questa domanda possiamo trovare una risposta mutuando un concetto dal linguaggio dei sociologi e della comunicazione: il *public engagement*, il coinvolgimento e la ricaduta della nostra collezione digitale sulla comunità di riferimento sarà il valore aggiunto,



Fiore di virtù, [Venezia], Beretin Convento Ca' Grande, 1477, carta a1r

soprattutto se si tratta di una raccolta molto di nicchia o di carattere locale come può essere, per esempio, la digitalizzazione di studi dedicati all'architettura e all'arte del territorio.

Una volta deciso cosa digitalizzare e come, e avendo pianificato le attività e magari attivato delle collaborazioni, per dare maggior impatto e visibilità a quanto si sta facendo, il passo successivo è cosa far vedere e soprattutto come farlo vedere. È dalle scelte che vengono fatte in questa fase che dipende la possibilità di far accedere gli utenti alle risorse digitalizzate. Non sempre è facile trovare risorse digitalizzate, e l'utente può trovarsi un po' spaesato di fronte a interfacce non sempre amichevoli come vorremmo che fossero, e a grandi quantità di risorse, difficili da individuare e selezionare.

Decidere come mostrare all'utente le informazioni relative alle risorse digitalizzate è qualcosa di intimamente collegato al lavoro "dietro le quinte" attuato in fase di catalogazione? Volendo sintetizzare al massimo, si può dire che se si mette in piedi uno scheletro solido e dettagliato di

metadati bibliografici, il più possibile granulari, dopo si può avere grande libertà ed enormi possibilità nel personalizzare la visualizzazione del catalogo, poiché si ha maggiore garanzia di efficacia nel recupero delle informazioni e nel raggruppamento dei record. Le realtà medio-piccole sono favorite in questo senso, poiché la grandezza del set di metadati da gestire non sarà mai sterminata e dunque la cura dei dettagli, la pulizia e la correttezza delle informazioni possono e devono diventare una cifra distintiva. Questo vale a maggior ragione nel caso in cui i record relativi ai libri digitalizzati confluiscono in un portale che accoglie anche risorse di diverso tipo. Per fare un esempio pratico, solo se ho adottato come norma l'utilizzo del titolo uniforme per identificare l'opera, sarò in grado di rintracciare, all'interno di un discovery tool, tutte le risorse digitali associate all'*Amleto* di William Shakespeare: la digitalizzazione di un'edizione settecentesca, quella presente in una raccolta delle opere dell'autore, la traccia audio-video di una messa in scena, l'audiolibro. Inoltre, più i metadati creati in origine sono precisi e accurati, maggiori saranno le possibilità di poterli riutilizzare anche al di fuori del proprio catalogo. Pensiamo a Wikidata: negli ultimi anni, sono stati sviluppati numerosi strumenti che permettono l'esportazione dei metadati di autorità dai cataloghi elettronici delle biblioteche, associando i singoli autori presenti nei record di partenza alle entità create dai wikipediani; questa attività, però, è molto più sicura se chi cataloga ha immesso informazioni precise per l'identificazione univoca dell'autore: ad esempio, la sola azione di inserire sempre le date di nascita e di morte, quando si crea un record di autorità relativo a una persona, abbassa esponenzialmente il grado di errore degli abbinamenti.

In una collezione digitale, strettamente connessi ai metadati bibliografici, troviamo i metadati strutturali: la loro funzione è quella di restituire le informazioni necessarie affinché una sequenza di immagini apparentemente scollegate diventi, o ritorni a essere, un oggetto con una sua struttura fisica e logica, nel nostro caso di un libro, composto quindi di introduzioni, capitoli, e paragrafi, che, ad esempio, inizia con un frontespizio e termina con un colophon.

Curare la redazione dei metadati strutturali in modo che siano coerenti e allineati a quelli bibliografici migliora l'esperienza di ricerca degli utenti, e aiuta a emulare l'esperienza della sala di lettura, per dirla con



Neu-eröffnetes Italien, Nürnberg, Johann Leonhard Buggel, 1703

le parole delle linee guida IFLA per la digitalizzazione di materiali rari e preziosi. Per fare degli esempi concreti, si pensi a quanto sia più efficace una ricerca in un periodico completamente spogliato o il valore aggiunto che porta a un catalogatore di libro antico un unico colpo d'occhio sulle tavole illustrate di un'opera o sugli errori di numerazione delle pagine; o ancora quanto sia più "naturale" leggere un testo digitale in Wikisource, arricchito di ripartizioni interne e con la possibilità di accedere direttamente ai capitoli, rispetto a tentare la stessa esperienza sul portale Internet Archive, dove i metadati sono scarsi e la lettura è sequenziale. Un progetto di indicizzazione di testi aperto e collaborativo come Wikisource può essere, inoltre, un valido esempio di come dei solidi metadati strutturali, con la loro organizzazione gerarchica, o ad albero, possano essere facilmente adattati per essere condivisi e valorizzati. Il funzionamento di Wikisource, infatti, prevede che i contenuti dei libri vengano descritti con un linguaggio di marcatura che li restituirà ai lettori secondo una precisa sequenza logica. La cura messa nella metadattazione, quindi, non servirà solo a rendere affidabili i nostri oggetti digitali, ma permetterà di poterli condividere e disseminare adattandoli facilmente ai contesti di destinazione che funzionano secondo la stessa logica, anche quando, come nel caso di Wikisource, non appartengono al mondo strettamente bibliotecario.



Pietro Andrea Mattioli, *Commentarii in sex libros Pedacii Dioscoridis Anazarbei de medica materia*, Venezia, Vincenzo Valgrisi, 1565

E infine, non si poteva non fare un accenno alla questione della conservazione del digitale. È certamente un argomento complesso, sfaccettato, che chiama in causa moltissimi fattori, per cui una sintesi è ardua, ma non per questo va trascurata o demandata a un secondo momento. Si tratta di attuare scelte strategiche importanti che avranno ricadute a lunga scadenza e che potranno portare a esiti negativi se non correttamente affrontate e valutate.

La conservazione del digitale è un argomento potenzialmente sterminato, soprattutto se si pensa alle risorse native digitali, come e-journals o ebooks: gli aspetti coinvolti in questo ambito sono molti, dal deposito legale digitale, alla conservazione intesa come garanzia di accesso perpetuo alle risorse cui la biblioteca si abbona, ai diritti commerciali associati e dunque ai vincoli con gli editori e i distributori. Tuttavia, se pensiamo alle digitalizzazioni di risorse che



Jacques Ozanam, *Nouveaux elements d'algebre*, Amsterdam, George Gallet, 1702, antiporta

nascono cartacee, i problemi relativi alla conservazione sono decisamente diversi. Nella maggior parte dei casi ciò che si digitalizza è materiale antico, svincolato dai diritti di copyright, pertanto la situazione tipica che una biblioteca si trova a fronteggiare è la seguente: si è avviato un progetto di digitalizzazione e sono disponibili online le risorse, ora bisogna gestire diversi Gigabyte di immagini. Qual è il modo giusto per conservarle, in modo tale che fra qualche anno non siano perdute o inutilizzabili? La prima raccomandazione da fare è di scegliere estensioni dei file immagine ad alta risoluzione che non temano l'obsolescenza e che non siano legate a software proprietari: i formati .TIFF o .RAW sono adatti a questo scopo. In secondo luogo, bisogna sempre ricordare che non basta conservare le

immagini, ma è indispensabile associarle ai metadati che le descrivono. Per fare questo, è necessario utilizzare identificativi univoci che permettano di rintracciare con sicurezza i dati e i metadati a essi collegati. Il terzo punto da affrontare riguarda lo spazio di archiviazione: i supporti di memorizzazione rimovibili, come gli hard disk esterni, possono essere utili durante la prima fase di gestione delle immagini, cioè per la loro acquisizione, il controllo di qualità e la lavorazione, ma non danno sicurezze nel lungo periodo. Di contro, una buona soluzione è quella di dotarsi di uno spazio di archiviazione in cloud che permetta il backup controllato dei dati: esistono molte società che offrono questo tipo di servizio a costi ragionevoli, con un elevato livello di sicurezza. Un'altra via percorribile, ma sicuramente molto più onerosa dal punto di vista economico, è quella dei sistemi di gestione integrati: in sostanza, più software che si interfacciano tra loro per gestire a tutto tondo le risorse, permettendone la catalogazione, la visualizzazione e la conservazione a lungo termine, quest'ultima attraverso l'utilizzo di repository pensati ad hoc.

Si ringrazia la fondazione BEIC per avere messo a disposizione il materiale iconografico che correda questo articolo.

NOTE

¹ Valeria De Francesca – Federica Viazzi, *Come gestire una collezione di libri digitalizzati*, Milano, Editrice Bibliografica, 2019.

ABSTRACT

The paper is a short reflection on the issue of creation, management and visualization of digitized book collection, starting from the reading of the book "Come gestire una collezione di libri digitalizzati" (How to manage a digitized book collection); the main topics, such as the steps of a digitization project, the issues of visualization and retrieval of the items, and the preservation of the collection, are discussed in an interview with the authors of the book.

DOI: 10.3302/0392-8586-201905-003-1