



Rivista del digitale nei beni culturali

ICCU-ROMA

Il Progetto di digitalizzazione dei Plutei della Biblioteca Medicea Laurenziana di Firenze*

Emiliano Degl'Innocenti

Società Internazionale per lo Studio del Medioevo Latino (SISMEL) di Firenze

La genesi

Un ambizioso progetto di digitalizzazione dei codici con segnatura “Plutei”, conservati presso la Biblioteca Medicea Laurenziana di Firenze, descritto nelle pagine seguenti, ha iniziato a prendere forma in seno al Comitato Guida per la Biblioteca Digitale Italiana già a partire dai primi mesi del 2005. Apparve già allora sufficientemente chiaro come nel nostro paese non avesse ancora attecchito lo stesso tipo di sensibilità che aveva portato, altrove in Europa – in particolare in Germania e in Svizzera – alla creazione di ampie collezioni di codici digitalizzati, liberamente fruibili in rete. Come è noto i fondi manoscritti presenti in Italia rappresentano una percentuale significativamente alta del patrimonio mondiale attestante la memoria delle civiltà del passato (in particolare greca, latina antica, medievale e rinascimentale); l'Italia riveste pertanto in Europa e nel mondo intero un ruolo di primo piano, sia per la conservazione, sia per la promozione della conoscenza di questa formidabile eredità culturale¹. A livello continentale, a fronte di un grande interesse per il tema delle *Digital Libraries* e della digitalizzazione del patrimonio culturale, risultavano già attivi – o erano in corso di attivazione – diversi importanti progetti² rivolti al mondo delle ri-

* Desidero ringraziare la Direttrice della Biblioteca, Franca Arduini, e le sue collaboratrici Ida Giovanna Rao e Sabina Magrini per l'importante contributo apportato alle diverse fasi evolutive del progetto, sommariamente descritte in queste pagine.

¹ Già nel 2003, nell'ambito della presidenza italiana del consiglio dell'Unione Europea il tema delle memorie digitali per la conservazione del patrimonio culturale fu di centrale importanza. Si ricordino a questo proposito, almeno i convegni *Futuro delle memorie digitali e patrimonio culturale* (http://www.imss.fi.it/memorie_digitali/indice.html) e *Archivi informatici per il patrimonio culturale* (<http://www.lincci.it/convegni/CONVEGNI.2004/PROGRAMMI/PRG.ARCHIVI.html>), che si svolsero nell'autunno di quell'anno. Gli atti di quest'ultimo sono recentemente stati pubblicati in volume: *Archivi informatici per il patrimonio culturale (Roma, 17-19 novembre 2003)*, Roma: Bardi, 2006. Tutti gli indirizzi citati nel presente articolo risultano attivi mentre scriviamo (maggio 2007), tuttavia qualora alcune delle risorse segnalate non fossero più raggiungibili direttamente, si potranno utilizzare gli *snapshot* presenti nella cache di Google, oppure i *mirror* archiviati dalla Wayback Machine di Internet Archive (<http://www.archive.org/web/web.php>).

² EDL (European Digital Library), <http://europa.eu.int/idabc/en/document/4239/330>, <http://www.edlproject.eu/>; TEL (The European Library), <http://www.theeuropeanlibrary.org/portal/index.htm>; Michael (Multilingual Inventory of Cultural Heritage in Europe), <http://www.michael-culture.org/it/home>.

sorse digitali e a particolari aspetti della gestione e disseminazione del sapere mediante strumenti tecnologici avanzati. In particolare, in diversi paesi europei, erano già nate interessanti iniziative di digitalizzazione di manoscritti – caratterizzate da differenti livelli di completezza, qualità e stato di avanzamento – fra le quali si segnalavano come esemplari Manuscriptorium³ (Biblioteca Nazionale di Praga); CESC, Codices Electronici Sangallenses⁴ (Stiftsbibliothek di S. Gallo); CEEC, Codices Electronici Ecclesiae Coloniensis⁵ (Dombibliothek di Colonia); Digitale Bibliothek-Münchener Digitalisierungszentrum⁶ (Bayerische Nationalbibliothek di Monaco di Baviera); Wolfenbütteler Digitale Bibliothek⁷ (Wolfenbüttel). Anche in Italia, nello stesso periodo, si registra la nascita di iniziative simili, come la digitalizzazione dei codici danteschi⁸, il Progetto Imerio⁹ o la digitalizzazione dei manoscritti appartenenti al Fondo Antico del Sacro Convento di Assisi¹⁰ (una collezione digitale di 709 mss., presto disponibile per la consultazione *on-line*).

A distanza di neppure un anno, nel 2006, l'urgenza di provvedere alla tutela del patrimonio manoscritto venne portata ancora più in primo piano delle allarmanti notizie provenienti dal Land tedesco del Baden-Württemberg, il cui governo aveva manifestato l'intenzione di porre in vendita, disperdendolo, l'importantissimo fondo manoscritto della Badische Landesbibliothek di Karlsruhe¹¹, suscitando le proteste della comunità scientifica internazionale.

L'idea che la digitalizzazione massiccia di fondi manoscritti – col duplice scopo di tutelare i supporti originali, «assicurare la trasmissione della civiltà [...] ai secoli futuri»¹² e fornire l'accesso a documenti non altrimenti consultabili¹³ (quali codici, mappe e manoscritti) – dovesse costituire una priorità ineludibile, fece sì che

³ http://www.manuscriptorium.com/Site/ENG/default_eng.asp.

⁴ <http://www.cesg.unifr.ch/it/index.htm>.

⁵ Gli aspetti più rilevanti del progetto sono descritti nel volume *Codices Electronici Ecclesiae Coloniensis. Eine mittelalterliche Kathedralbibliothek in digitaler Form*, hg. Manfred Thaller, Göttingen: Duehrkohp & Radicke, 2001. La biblioteca digitale vera e propria è raggiungibile all'indirizzo <http://www.ceec.uni-koeln.de/>.

⁶ <http://mdz1.bib-bvb.de/~mdz/>.

⁷ <http://www.hab.de/bibliothek/wdb/index.htm>.

⁸ Progetto che si è avvalso della consulenza scientifica della Società Dantesca Italiana (<http://www.danteonline.it>).

⁹ Progetto di riproduzione e classificazione su supporto digitale dei codici giuridici e filosofico-teologici del Collegio di Spagna in Bologna (<http://www.cirsfid.unibo.it/imerio/presenta.php>).

¹⁰ Sostenuta dal Ministero per i Beni e le Attività Culturali, Direzione Generale per i Beni Librari e gli Istituti Culturali (<http://www.sisf-assisi.it/digitalizzazione.htm>).

¹¹ <http://archiv.twoday.net/stories/2743873/>; <http://www.blb-karlsruhe.de/blb/images/2006/presse-medioevolatino.pdf>.

¹² Cfr. Tullio Gregory, *Origini della terminologia filosofica moderna. Linee di Ricerca*, Firenze: Olschki, 2006, p. 3-32 (la citazione è estrapolata dalla p. 3).

¹³ Questo punto costituisce una importante peculiarità delle biblioteche digitali di manoscritti. A differenza dei prodotti seriali della stampa moderna, ciascun codice rappresenta un *unicum* irripetibile: se la perdita di un raro esemplare a stampa può essere considerata grave o gravissima, quella di un testimone manoscritto può rivelarsi davvero irreparabile.

Claudio Leonardi – Accademico dei Lincei e presidente della Società Internazionale per lo Studio del Medioevo Latino (SISMEL) – sollecitato e sostenuto dal Comitato Guida per la Biblioteca Digitale Italiana, elaborasse un articolato progetto di digitalizzazione dei manoscritti conservati presso diverse biblioteche italiane¹⁴, che tuttavia – per varie ragioni – non risultò immediatamente attuabile nella sua interezza. Ma ormai, la necessità indifferibile di metter mano ad un ampio progetto di trasferimento in ambiente digitale dell'identità culturale italiana, non poteva più essere accantonata: grazie a questa nuova sensibilità, e al concreto impegno della Direzione Generale per i Beni Librari e dell'Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), fu deciso a questo riguardo di varare un ambizioso piano di digitalizzazione riguardante le differenti tipologie di quello che venne definito il "libro italiano": dal manoscritto alla stampa antica, dalla cultura medievale e umanistica alla scienza moderna.

Per l'ambito specifico che qui intendiamo trattare, relativo alla cultura manoscritta, dalla prima stesura del progetto elaborato da Claudio Leonardi fu estrapolata, a titolo di esperienza pilota, sia per il prestigio dell'istituzione, che per l'assoluta rilevanza del fondo manoscritto, la collezione della Biblioteca Medicea Laurenziana¹⁵ di Firenze. Il piano iniziale fu rimodulato per far fronte ai mutati obiettivi e, dopo una prima fase di studio – condotta in stretta collaborazione con la stessa Biblioteca fiorentina – la sua Direttrice, persuasa della fattibilità dell'intervento e consapevole dell'importanza del progetto, dette il proprio assenso, siglandone di fatto l'atto di nascita ufficiale.

La prima fase di vita del progetto di digitalizzazione dei Plutei della Biblioteca Medicea Laurenziana è consistita nell'espletamento di una serie di importanti e complesse operazioni preparatorie, che hanno portato alla redazione di uno studio di fattibilità e alla effettiva scrittura della documentazione tecnico-amministrativa necessaria: parallelamente, da parte di gruppi di lavoro composti da esperti e rappresentanti delle istituzioni coinvolte (SISMEL¹⁶, Biblioteca Medicea Laurenziana¹⁷ e ICCU¹⁸), sono stati approfonditi diversi temi: dall'utilizzo del set di metadati MAG (Metadati Amministrativi Gestionali)¹⁹, ad aspetti legati alle dinamiche di gestione

¹⁴ Si trattava in origine di 11 sedi di conservazione, dislocate su tutto il territorio nazionale.

¹⁵ Nelle motivazioni addotte in fase di presentazione del progetto ci si esprimeva in questo modo: «la decisione di iniziare tale lavoro dalla Biblioteca Medicea Laurenziana di Firenze si giustifica per lo straordinario patrimonio manoscritto posseduto dalla Biblioteca, come anche per la fama e la notorietà di molti dei suoi manoscritti».

¹⁶ Alla Società Internazionale per lo Studio del Medioevo Latino di Firenze è stato affidato il ruolo di coordinamento tecnico-scientifico dell'intero progetto (<http://www.sismelfirenze.it>).

¹⁷ <http://www.bml.firenze.sbn.it>.

¹⁸ <http://www.iccu.sbn.it/genera.jsp>.

¹⁹ Il set MAG (<http://www.iccu.sbn.it/genera.jsp?id=105>) è stato utilizzato per la codifica dei metadati di base relativi ai singoli mss. (autore, titolo dell'opera, opera anonima, datazione e segnatura) desunti dai cataloghi a stampa di Bandini Assemani e Biscioni. L'argomento è discusso anche in seguito.

delle risorse digitali prodotte durante il ciclo di vita del progetto (dall'interoperabilità alla *long term preservation*). Attualmente, conclusa questa fase, sono in corso di espletamento le operazioni di valutazione delle proposte elaborate dai partecipanti alla gara d'appalto bandita per la fornitura dei servizi di digitalizzazione e indicizzazione del fondo manoscritto²⁰, e vi sono fondate speranze che i lavori possano avere effettivamente inizio in tempi piuttosto brevi, per concludersi nell'arco di 30 mesi.

Il fondo²¹

Come è già stato evidenziato, il criterio guida che ha condotto alla scelta del fondo manoscritto della Biblioteca Medicea Laurenziana è stato quello dell'assoluta eccellenza: la raccolta dei manoscritti posseduti dalla Biblioteca fiorentina rappresenta infatti – nel pur relevantissimo patrimonio del paese – un caso unico e di straordinaria importanza, sia perché conserva alcuni dei più importanti, rari e antichi testimoni della produzione culturale occidentale, sia perché raccoglie un numero rilevante di codici che documentano la produzione libraria entro un arco cronologico che va dal V al XIX sec. Fra i manoscritti conservati presso la Biblioteca Medicea Laurenziana si annoverano autori come Tacito, Plinio, Eschilo, Sofocle, Quintiliano e Virgilio, oltre a preziosi codici autografi di Petrarca e Boccaccio. La collezione inoltre testimonia alcuni dei momenti più importanti della storia del Rinascimento fiorentino, dalla sua nascita alla maturità, documentando l'Umanesimo con la presenza (come autori, copisti e possessori) di figure di prima grandezza come Coluccio Salutati, Poggio Bracciolini, Niccolò Niccoli, Marsilio Ficino e Pico della Mirandola. In essa anche le principali correnti artistiche del momento risultano ben rappresentate da splendidi codici miniati da maestri come Attavante, Francesco di Antonio del Chierico, Francesco Rosselli o i fratelli Del Fora.

Sebbene il nucleo originario di codici²² sia costituito dalla raccolta privata della famiglia Medici, iniziata da Cosimo di Giovanni di Bicci (1389-1464), la collezione è andata via via incrementandosi, sia mediante l'acquisizione di intere biblioteche di importanti umanisti come Francesco Sasseti (1421-1490) e Francesco Filelfo (1398-1481), sia grazie all'arrivo di vari manoscritti provenienti dal convento domenicano di San Marco.

I "plutei puri", di provenienza privata medicea, assommano ben 3007 volumi e debbono il loro nome agli 88 banchi di noce – appositamente disegnati da Michelangelo per accoglierli e ordinali, detti appunto "plutei" – su cui risultavano disposti nel 1571, in occasione della prima apertura al pubblico della biblioteca,

²⁰ Per un importo totale di 1.250.000 euro.

²¹ Per le informazioni riguardanti la storia della Biblioteca Medicea Laurenziana desidero ringraziare la Dr.ssa Ida Giovanna Rao, responsabile dell'Ufficio manoscritti, papiri, tutela monumentale della Biblioteca fiorentina, che ha avuto la gentilezza di fornirne. Va da sé che la responsabilità per eventuali errori o imprecisioni in esse presenti, dovrà essere attribuita unicamente a chi scrive.

²² I cosiddetti "plutei puri", che – assieme ai plutei "superiore", "inferiore", "sinistra" e "destra" – costituiscono i fondi più antichi della Biblioteca, oggetto della campagna di digitalizzazione.

coincidente con il compleanno del granduca Cosimo I.

I “plutei superiore e inferiore”, che raccolgono circa 403 volumi, sono costituiti da quella che fu la biblioteca della importante famiglia fiorentina dai Gaddi, giunta nel 1755 a rimpinguare la collezione medicea, dalla raccolta del canonico Anton Maria Biscioni (1674-1756), arrivata nel 1756, e da 14 codici dei frati riformati di Montepulciano, pervenuti nel 1758: tutti i volumi di nuova acquisizione furono collocati in scaffali aggiuntivi rispetto agli 88 banchi originari, numerati dall’89 al 91 e denominati sempre “plutei”, ma con l’aggiunta di “sup.” e “inf.”.

Altri 595 manoscritti, denominati “plutei sinistra e destra” appartennero originariamente alla libreria di Santa Croce, e furono destinati alla Biblioteca Medicea Laurenziana nel 1767: essi vennero disposti esattamente seguendo la loro collocazione originale, che era di 36 banchi “*ex parte orientis*” e 36 “*ex parte occidentis*”. Anche in questo caso, come nei precedenti, i volumi vennero denominati “plutei”, stavolta con la variante “sinistra” e “destra”.

La biblioteca digitale

Se appare immediatamente evidente il valore che un simile intervento può rivestire per la comunità scientifica (dal paleografo al linguista, dal filosofo al filologo, e non solo), non sembra tuttavia superfluo evidenziare come, attraverso la messa in rete delle riproduzioni integrali di questi codici, il progetto di digitalizzazione dei Plutei della Biblioteca Medicea Laurenziana intenda promuovere la conoscenza di uno dei più importanti fondi manoscritti del mondo, anche fra i non addetti ai lavori, e favorire l’accesso ad un patrimonio culturale di inestimabile valore anche a tutti gli utenti che non sono in grado – per i motivi più diversi – di recarsi fisicamente negli ambienti michelangioleschi della biblioteca fiorentina, esperienza questa che resta, peraltro, vivamente consigliata²³.

La collezione digitale risultante dalle operazioni di acquisizione numerica dei manoscritti e codifica dei metadati sarà costituita da oltre 1.350.000 immagini, corrispondenti a più di 3.900 manoscritti integralmente riprodotti, fra cui si annoverano, come abbiamo già segnalato, alcuni dei più rilevanti testimoni della produzione culturale e intellettuale dell’Occidente (e non solo, dal momento che la biblioteca conserva anche codici orientali di grande valore). Queste immagini saranno corredate dalle informazioni di carattere scientifico, provenienti dal recupero – in formato digitale – dei tre principali cataloghi a stampa settecenteschi²⁴ che descrivono i codici appartenenti al fondo.

²³ Anche se, sul portale Internet Culturale è possibile visitare la biblioteca mediante un *tour* virtuale (<http://www.internetculturale.it/genera.jsp?id=813>).

²⁴ Approntati da Stefano Evodio Assemani, erudito e arcivescovo di Apamea, *scriptor* di siriano e arabo e custode principale alla Biblioteca Apostolica Vaticana (1742), dal canonico Anton Maria Biscioni, prefetto della Laurenziana e esperto orientista (1752) e dal suo successore Angelo Maria Bandini (1764-1793), al quale si deve in effetti il primo grande catalogo di tutto il posseduto della libreria Medicea (esclusi i codici orientali).

Il progetto, è già stato sottolineato, si pone come obiettivi primari la valorizzazione della nostra eredità culturale in ambiente digitale, mediante la disseminazione delle informazioni e la veicolazione dei contenuti ad un numero di soggetti che non sia limitato unicamente agli specialisti di settore, ma che – mediante l’implementazione una serie di strumenti e servizi innovativi – sia in grado di soddisfare fasce di utenza diversificate, e la conservazione sul lungo periodo delle risorse digitali prodotte, nell’ottica del potenziamento degli strumenti elettronici per la tutela e la diffusione di un patrimonio di inestimabile valore e ricchezza. Tali obiettivi dovranno essere raggiunti – in continuità con le linee guida operative e gli standard adottati dalla Biblioteca Digitale Italiana – mediante la costituzione di una collezione digitale di immagini e metadati, la messa a punto di un ambiente di fruizione efficiente, amichevole e dotato di efficaci strumenti di *information retrieval*, l’implementazione di uno spazio di archiviazione affidabile (*trusted digital repository*) per la conservazione degli oggetti digitali, l’adozione di tecnologie adeguate, standard aperti e documentati per il loro trattamento, e l’utilizzo di opportune strategie di codifica, gestione e indicizzazione dei dati per facilitare e rendere proficua l’esperienza degli utenti.

Le fasi del progetto

Dal punto di vista operativo il progetto si articolerà in tre distinte fasi: la scansione delle immagini, la produzione dei metadati ad esse relativi e il caricamento degli oggetti digitali sulla teca della Biblioteca Medicea Laurenziana. Le prime due, ovvero l’acquisizione in formato digitale degli originali (manoscritti e cataloghi a stampa) e la produzione dei relativi metadati, necessari per l’indicizzazione del materiale acquisito, verranno portate avanti parallelamente, in modo da ottimizzare i flussi di lavoro. La terza e ultima fase prevede il caricamento vero e proprio delle immagini e dei metadati prodotti sulla teca digitale della Biblioteca Medicea Laurenziana: lungi dal risultare una mera formalità, essa è invece di fondamentale importanza, in quanto consentirà ai responsabili del progetto di verificare sul campo – ad intervalli prefissati e regolari²⁵ – il livello qualitativo effettivamente raggiunto (e dunque la bontà o meno delle scelte intraprese), sia per quanto concerne le caratteristiche delle immagini e dei metadati, che per tutti gli aspetti legati alla qualità del servizio erogato, alla accessibilità dei contenuti e alla utilizzabilità dell’interfaccia utente.

Digitalizzazione degli originali

Durante la prima fase, i supporti originali (cartacei e membranacei), verranno sottoposti a processi di scansione per l’acquisizione in formato digitale: appare persino superfluo segnalare come, in considerazione dello straordinario valore dei codi-

²⁵ Secondo le scansioni previste dal cronoprogramma allegato alla documentazione di gara, disponibile sul sito della Biblioteca Medicea Laurenziana (http://www.bml.firenze.sbn.it/bandi_di_gara.htm).

ci Laurenziani, le apparecchiature utilizzate dovranno garantire oltre alle migliori prestazioni (sia dal punto di vista della qualità delle immagini risultanti, che da quello della velocità di esecuzione), il più alto livello di tutela dell'integrità dei supporti scrittori. In estrema sintesi, gli apparati di ripresa utilizzati dovranno essere caratterizzati dalla totale assenza di impatto potenzialmente negativo sui manoscritti e sugli stampati, risultare in grado di acquisire immagini ad altissima risoluzione, e garantire una alta velocità d'acquisizione e trasferimento dei dati alle postazioni di post-elaborazione, controllo e memorizzazione, per non generare rallentamenti e colli di bottiglia nei flussi di lavoro.

Durante questa prima e delicata fase del progetto verranno prodotti, per ciascuna immagine acquisita, due differenti tipi di file, caratterizzati da standard qualitativi differenti e diverse destinazioni d'uso. Ciascuna immagine sarà archiviata sia in un formato adeguato alla conservazione a lungo termine, denominato *master*, che costituirà una sorta di surrogato digitale dell'originale e che quindi dovrà garantire un livello di qualità eccellente, sia in un formato più leggero e maneggevole, qualitativamente più modesto rispetto al precedente, destinato alla consultazione in linea, comunemente definito "formato *internet*". L'elevata risoluzione delle immagini *master* consentirà di limitare, laddove necessario, il ricorso diretto agli originali (favorendone così la tutela e la conservazione), e disporre di un'alternativa qualitativamente superiore rispetto ai microfilm esistenti, consentendo – ad esempio – di fornire agli utenti che ne facessero richiesta riproduzioni digitali di altissima qualità del materiale originale.

Dal punto di vista tecnico, le caratteristiche delle immagini *master* risultano ampiamente in linea con quelle utilizzate, in ambito internazionale, da altri importanti progetti di digitalizzazione: risoluzione di 600 dpi (*dot per inch*, punti per pollice²⁶) non interpolati²⁷, formato di codifica dei file TIFF²⁸ non compresso, con una profondità del colore di 24 bit RGB²⁹. L'alta qualità delle immagini così prodotte è testimoniata, in maniera diretta, dal peso dei file risultanti: da una stima effettuata durante la fase preparatoria del progetto, il peso medio di una immagine *master* si attesterà intorno ai 70 MB (con variazioni, anche significative, dipendenti dalle dimensioni dei supporti originali). Questo significa che, teoricamente, per conservare l'intera collezione digitale della Biblioteca Laurenziana sarebbero necessari oltre 100 *Terabyte*³⁰ di spazio su disco.

²⁶ Per una definizione cfr.: http://en.wikipedia.org/wiki/Dots_per_inch.

²⁷ Si tratta di un processo di elaborazione delle immagini utilizzato per aumentare in maniera artificiale la quantità di informazioni raccolte dal sistema di scansione.

²⁸ Cfr. la voce corrispondente in Wikipedia: http://it.wikipedia.org/wiki/Tagged_Image_File_Format.

²⁹ Cfr. http://it.wikipedia.org/wiki/Profondit%C3%A0_di_colore.

³⁰ Il termine è utilizzato in questo contesto per esprimere unicamente l'ordine di grandezza "mille miliardi di byte" (10¹² byte). Cfr. la voce corrispondente in Wikipedia: <http://it.wikipedia.org/wiki/Terabyte>.

Fileformat	TIFF (Tagged Image File Format) 6.0, non compresso
Risoluzione	600 dpi (<i>dot per inch</i> , punti per pollice) ottici
Profondità del colore	24 bit RGB (Truecolor)
Peso medio per ciascuna immagine	intorno ai 70 MB
Spazio su disco occupato (<i>off-line</i> su LTO)	circa 100 Terabyte

Anche per quanto riguarda le specifiche tecniche relative ai file destinati alla consultazione in linea ci troviamo di fronte a caratteristiche standard, che risultano in sintonia con quelle previste da altri progetti nazionali e internazionali: la qualità delle immagini, acquisite alla risoluzione di 72 dpi e archiviate in formato JPEG³¹ compresso, con una profondità del colore di 24 bit RGB, permetterà ovviamente la piena leggibilità e fruibilità dei contenuti (per motivi di studio e consultazione), ma non consentirà la loro riproduzione per scopi commerciali.

Anche in considerazione della limitata penetrazione della banda larga (ADSL e connessioni veloci) nel nostro paese – confermata dai dati forniti da Eurobarometro – che si attesterebbe intorno al 14% a fronte del doppio circa della media europea³², si è ritenuto opportuno contenere il peso delle immagini destinate alla fruizione in rete entro gli 1,5 MB, così da permettere ad un più ampio numero di utenti di fruire in maniera soddisfacente del servizio offerto. Lo spazio su disco richiesto per l'archiviazione e il mantenimento *on-line* della collezione digitale si attesta attorno ai 2,5 Terabyte, un valore che fino a qualche anno fa poteva sembrare improponibile, ma che – allo stato attuale della tecnologia – appare del tutto adeguato e sostenibile.

Fileformat	JPEG (Joint Photographic Experts Group) compresso
Risoluzione	72 dpi (<i>dot per inch</i> , punti per pollice) ottici
Profondità del colore	24 bit RGB (Truecolor)
Peso medio per ciascuna immagine	intorno ad 1,5 MB
Spazio su disco occupato (<i>on-line</i>)	circa 2,5 Terabyte

³¹ Cfr. Wikipedia: http://it.wikipedia.org/wiki/Joint_Photographic_Experts_Group.

³² Cfr. l'Eurobarometro pubblicato nell'aprile di quest'anno, disponibile all'indirizzo http://ec.europa.eu/information_society/policy/ecommm/doc/info_centre/studies_ext_consult/ecommm_household_study/eb07_finalreport_v4.pdf, in cui si evidenzia il netto ritardo del nostro paese nell'accesso alla banda larga. Neppure il tasso di crescita, stimato secondo lo stesso studio attorno al 3%, appare particolarmente incoraggiante, specie se confrontato con quello di altri paesi europei, in particolare Regno Unito (+9%), Germania e Francia (+6%).

Codifica dei metadati e strumenti di ricerca

Sin dalle primissime fasi di sviluppo del progetto è emersa con chiarezza la necessità di definire con precisione quali fossero le informazioni di carattere scientifico su ciascun manoscritto da fornire agli utenti di una così ricca collezione digitale, e di precisarne il formato di codifica. Dopo una approfondita analisi della questione, il Comitato Guida – in considerazione degli ambiziosi obiettivi del progetto e compatibilmente con i tempi previsti per il suo completamento – ritenne opportuno adottare una strategia piuttosto pragmatica, che non prevedesse una catalogazione *ex-novo* dei manoscritti appartenenti al fondo da digitalizzare, ma recuperasse il nucleo di metadati ritenuti indispensabili direttamente dai cataloghi a stampa esistenti. Ovviamente il formato prescelto per la loro codifica fu MAG³³, un profilo applicativo XML (eXtensible Markup Language) per la codifica dei Metadati Amministrativi e Gestionali promosso dall'ICCU e adottato dalla totalità dei progetti finanziati dal Comitato Guida per la Biblioteca Digitale Italiana.

Per consentire agli utenti una più agevole fruizione delle risorse digitali messe a disposizione dal progetto, il contenuto dei singoli codici verrà indicizzato a partire dalle informazioni contenute nei cataloghi a stampa, approntati dagli eruditi settecenteschi Assemani³⁴, Biscioni³⁵ e Bandini³⁶: in particolare, verranno recuperate, opportunamente trattate e integrate nei file XML-MAG, le signature dei manoscritti, oltre ad informazioni relative alla loro datazione e al loro contenuto, come i nomi degli autori e i titoli delle loro opere, oppure i titoli delle opere anonime:

- signature: indispensabile per consentire l'esatta individuazione di ciascun codice, può essere considerata la chiave primaria per l'accesso al materiale digitalizzato;
- datazione del codice: per il trattamento di questo dato si prevede l'utilizzo di etichette descrittive (es. *XIII sec.*) e forchette cronologiche espresse mediante valori numerici (es. *1201-1300*), raccordate fra loro mediante l'utilizzo di opportune tabelle di conversione (es. *XIII sec. = 1201-1300*);
- autore dell'opera (es. *Vergilius poeta*);
- titolo dell'opera (es. *Bucolica*).

³³ Cfr. la nota 19.

³⁴ Stefano Evodio Assemani, *Bibliothecae Mediceae Laurentianae et Palatinae codicum mms. Orientalium Catalogus*, Florentiae: Ex typographio Albiziniano, 1742.

³⁵ Anton Maria Biscioni, *Bibliothecae Mediceo-Laurentianae Catalogus*, Florentiae: ex Imperiali Typographio, 1752.

³⁶ Angelo Maria Bandini, *Catalogus codicum manuscriptorum Bibliothecae Mediceae Laurentianae varia continens opera graecorum patrum*, I-III, Florentiae: Typis Caesareis, 1764-1770; Angelo Maria Bandini, *Catalogus codicum latinorum Bibliothecae Mediceae Laurentianae sub auspiciis Petri Leopoldi*, I-V, Florentiae: s.n., 1774-1778; Angelo Maria Bandini, *Bibliotheca Leopoldina Laurentiana, seu, Catalogus manuscriptorum qui iussu Petri Leopoldi*, I-III, Florentiae: Typis Caesareis, 1791-1793.

Il sistema consentirà di effettuare ricerche complesse nell'intero *repository*, anche incrociando i vari campi (es. *Vergilius + XIII sec.*). A questo primo strato di metadati descrittivi potranno essere in seguito essere aggiunte altre informazioni, come:

- numero delle carte per ciascun ms.;
- tipologia del ms. (es. *cartaceo* [cart.] o *membranaceo* [membr.]);
- eventuali altri dati relativi alla presenza di elementi di pregio, quali miniature, ecc.

Le informazioni raccolte verranno ospitate nella sezione bib dello schema MAG:

- `<dc:creator>`: serve per indicare il nome dell'autore della risorsa. Nel nostro caso potrà contenere i nomi degli autori presenti all'interno del ms., come indicato da Bandini, Assemani e Biscioni:
Es.: `<dc:creator>Aristoteles</dc:creator>`;
- `<dc:title>`: indica il titolo della risorsa, o, più formalmente, il nome attraverso il quale la risorsa è conosciuta. Nel nostro caso potrà contenere i titoli delle opere (di autori conosciuti oppure anonime) presenti nei mss., secondo le indicazioni fornite dai cataloghi di Bandini, Assemani e Biscioni:
Es.: `<dc:title>Carmina XXVII</dc.title>`;
- `<dc:date>`: esprime una data associata a un evento del ciclo di vita della risorsa. Nel nostro caso può contenere – sotto forma di espressione di tipo testuale – la datazione del singolo ms., con indicazione di secolo e parte (es. *XII sec. prima metà*), corredata dall'indicazione della forchetta cronologica corrispondente (es.: *1101-1149*), espressa mediante valori numerici;
- `<holdings>`: serve a localizzare l'oggetto analogico all'interno di una particolare istituzione. Nel nostro caso, assieme agli elementi opzionali:
 - `<library>`, che contiene il nome dell'istituzione proprietaria dell'oggetto analogico o di parte dell'oggetto analogico;
 - `<shelfmark>`, che contiene la collocazione dell'oggetto analogico all'interno del catalogo dell'istituzione che lo possiede, potrà essere utilizzato per immagazzinare i dati relativi alla segnatura, mediante cui si è soliti indicare ciascun manoscritto all'interno dei cataloghi a stampa.

Gli utenti interessati potranno inoltre attingere ulteriori informazioni circa la bibliografia recente relativa a ciascun codice digitalizzato, raccolte nel corso degli anni da studiosi e personale della Biblioteca Medicea Laurenziana e conservate in una apposita banca dati specializzata, che verrà integrata nel sistema. Per favorire l'interoperabilità e il dialogo con altri sistemi e consentire l'accesso ai contenuti della collezione digitale anche da parte degli utenti del portale BDI è stato scelto come protocollo di esposizione dei metadati OAI-PMH (Open

Archives Initiative, Protocol for Metadata Harvesting)³⁷, lo standard *de facto* per l'implementazione di servizi a valore aggiunto nell'ambito della biblioteca digitale.

L'ambiente di fruizione

Grazie alle modalità di diffusione delle informazioni sopra descritte e alle procedure di *metadata harvesting* (collezione dei metadati) messe in atto dall'ICCU, l'utente – attraverso una interfaccia web semplice e consistente – potrà localizzare e quindi accedere agli oggetti digitali di proprio interesse, sia collegandosi direttamente al sito istituzionale della Biblioteca Medicea Laurenziana, sia attraverso i servizi a valore aggiunto implementati dal portale della Biblioteca Digitale Italiana. Utilizzando le possibilità di interrogazione messe a disposizione dal sistema sarà possibile navigare all'interno della collezione digitale utilizzando le categorie sopra elencate (autore, titolo dell'opera, opera anonima, datazione e segnatura) come chiavi di ricerca e – attraverso gli strumenti di cui è dotato il modulo di fruizione – visualizzare le riproduzioni digitali integrali delle schede redatte da Assemani, Bandini e Biscioni (relative ai singoli manoscritti) o sfogliare le pagine dei codici, proprio come si farebbe trovandosi in biblioteca.

La sostenibilità del progetto

Dal punto di vista tecnologico, anche in considerazione dei processi di rapida obsolescenza a cui sono soggette le infrastrutture hardware e software – spesso imposti dalle fluttuazioni del mercato più che dall'effettivo susseguirsi dei cicli tecnologici – si è ritenuto opportuno utilizzare per la gestione del progetto standard tecnologie aperte e ben documentate, limitando il più possibile il ricorso a soluzioni proprietarie, in modo da garantire un buon livello di sostenibilità (sia dal punto di vista tecnico che da quello economico) al progetto stesso. È stata inoltre riconosciuta l'importanza di valutare per tempo l'adozione di opportune strategie per preservare nel tempo l'integrità e l'accessibilità delle risorse digitali prodotte e – in ultima analisi – garantire la continuità e la qualità del servizio erogato. Sono state a questo proposito valutate diverse soluzioni, dalla migrazione dei dati (rispetto, ad esempio all'evoluzione dei sistemi di *storage*, dei formati di codifica e dei sistemi operativi utilizzati), alla previsione di periodici *refresh* delle collezioni digitali (immagini e metadati), al fine di garantire la fruibilità nel tempo, alla ridondanza e differenziazione dei supporti utilizzati per la memorizzazione delle informazioni, all'adozione di efficaci strategie di lungo periodo per la manutenzione della teca digitale che ospiterà le immagini e i metadati, ecc.

³⁷ Cfr. la voce corrispondente in Wikipedia: http://it.wikipedia.org/wiki/Open_Archives_Initiative.

Conclusioni

Essendomi già soffermato a più riprese sugli aspetti qualitativi e tecnologici del progetto, in chiusura di questa breve rassegna ritengo utile fornire alcuni essenziali dati numerici, per favorire la comprensione della sua reale dimensione quantitativa, anche in relazione ad altre iniziative simili citate in precedenza: la collezione digitale della Biblioteca Medicea Laurenziana, una volta completata, metterà a disposizione degli utenti un numero di codici dalle 4,5 alle 100 volte superiore rispetto ai progetti ricordati all'inizio³⁸, candidandola al ruolo di risorsa chiave per la diffusione e la traduzione della nostra eredità culturale in ambiente digitale, ben consapevoli che «ogni cultura è continua trascrizione, traduzione di precedenti esperienze che, trasferite in altri contesti, portano nuovi frutti e permettono di scoprire nuovi orizzonti del sapere»³⁹.

³⁸ In particolare, utilizzando i dati diffusi sui rispettivi siti web, aggiornati al maggio del 2007 (fra parentesi il numero di codici interamente digitalizzati): *CESG - Codices Electronici Sangallenses* (131 mss.); *CEEC - Codices Electronici Ecclesiae Coloniensis* (320 mss.); *Wolfenbütteler Digitale Bibliothek* (82 mss.); *Dante on-line* (33 mss.); *Progetto Irnerio* (297 mss.); Fondo Antico del Sacro Convento di Assisi (709 mss.).

³⁹ Sperone Speroni, *Dialogue des langues*, édition bilingue, traduction de Gérard Genot et Paul Larivaille, introduction et notes de Mario Pozzi, texte établie par Mario Pozzi, Paris: Les Belles Lettres, 2001 (citato in Tullio Gregory, *Origini della terminologia filosofica moderna* cit., p. 35).