

Livelli descrittivi, relazioni e contesti di produzione nella Sapienza Digital Library

Cecilia Carloni¹ Maria Guercio¹

¹ DigiLab Sapienza, Roma, Italia
{cecilia.carloni,maria.guercio}@uniroma1.it

Abstract. L'ambiente informativo della Sapienza Digital Library¹ intende raccogliere e rendere accessibili materiali di diversa natura e provenienza preservando i contesti di produzione, la profondità informativa e la complessità semantica dei patrimoni documentari. Una continua ricerca e sperimentazione degli strumenti tecnici e concettuali mira a costruire un'architettura dei contenuti che espliciti il contesto di produzione e le relazioni strutturali e semantiche, prevedendo quindi una descrizione a più livelli dei complessi documentari. Solo nuove forme di intermediazione tra sistemi informativi complessi e utenti possono veicolare le tracce culturali e i contesti che le hanno prodotte. Il modello di rappresentazione dei contenuti informativi distingue il piano dell'oggetto fisico da quello della rappresentazione digitale e semantica.

Parole chiave: biblioteche digitali, archivi digitali, collezioni, integrazione, convergenza digitale, interoperabilità, standard descrittivi.

1 Introduzione

Il progetto Sapienza Digital Library (d'ora in avanti SDL) nasce con la missione di aggregare e rendere accessibile la documentazione di diversa natura e provenienza prodotta sia dai gruppi di ricerca operanti all'interno dell'Università, sia da enti, famiglie e persone esterne alla comunità accademica. Si intende quindi sviluppare un'architettura informativa che possa accogliere materiali diversi per tipologia (libri, periodici, spartiti musicali, fotografie, materiale multimediale, film, filmati audio e video, registrazioni sonore, ricostruzioni 3D di luoghi, edifici e territori) e provenienza, in quanto riconducibili a diverse modalità di produzione e sedimentazione di fonti documentarie e del patrimonio culturale -archivi, biblioteche, musei- nonché a differenti domini metodologici e disciplinari.

Il portale è inoltre progettato per consentire l'interrogazione congiunta, dalla medesima interfaccia, dei cataloghi delle biblioteche del Sistema bibliotecario di Ateneo, delle riviste elettroniche sottoscritte ad abbonamento e dei prodotti della ricerca caricati su U-Gov, oltre che – naturalmente – delle diverse fonti digitali acquisite e catalogate/inventariate all'interno della SDL e delle risorse digitali di provenienza esterna,

¹ Il prototipo sarà accessibile a maggio 2013. URL= <http://sdl.uniroma1.it>

donate da utenti esterni o in quanto materiali di loro produzione o quali perché ne detengono i diritti d'uso.

All'interno della nostra comunità universitaria, come in molte altre istituzioni in ambito culturale e scientifico, esistono e vengono continuamente prodotti numerosi e importanti archivi e collezioni di diversa natura, ma senza dubbio rilevanti sul piano culturale e scientifico: un vasto universo documentario in forma analogica, digitalizzata o nativa digitale relativo a opere d'arte, monumenti e collezioni, archivi di ricerca, aree di scavo e siti preistorici, data set, musei virtuali (Capaldi, Iardi, Ragone 2011; Ragone 2012). Solo alcuni di questi contenuti digitali sono accessibili sul web, in molti casi descritti in maniera piatta e non controllata e pubblicati su piattaforme diverse non integrate da un unico motore di ricerca. In altri casi, si tratta di materiale conservato in forma analogica, per lo più chiuso in armadi dimenticati e solo raramente, e non sempre correttamente, reso disponibile all'interno della stessa comunità disciplinare di cui il produttore o detentore del patrimonio pur è parte.

L'obiettivo prioritario di SDL è quello di integrare e rendere visibili attraverso un accesso unificato collezioni e archivi finora dimenticati o poco valorizzati, supportando strutture dell'Università e soggetti esterni che vogliono comunicare i propri patrimoni attraverso il portale SDL, in molti casi rivitalizzando mediante l'accessibilità sul web patrimoni informativi e di conoscenza specialistici sedimentati (ma non ordinati né valorizzati) nel corso di attività di ricerca, di cui tradizionalmente ci si limita a pubblicare i risultati principali. Le istituzioni del mondo culturale hanno un ruolo fondamentale nell'assicurare la rilevanza scientifica dei contenuti digitali e nel mediare tra i complessi documentari e gli utenti e garantiscono la qualità di tale processo anche mediante lo sviluppo di strumenti di ricerca appropriati in grado di salvaguardare il contesto documentario e di produzione e apparati descrittivi controllati, di qualità scientifica e filologicamente rigorosi.

La trasposizione in digitale non deve alterare la natura organica e sistemica degli archivi e delle collezioni, né impoverirne il contenuto informativo e le relazioni, ma deve piuttosto tendere a fornire qualità di contenuti e affidabilità. Il web offre nuovi strumenti di recupero e condivisione e consente di arricchire le risorse digitali o digitalizzate con informazioni di rappresentazione che descrivono il fitto tessuto di relazioni verticali e trasversali interne alle collezioni/archivi o di collegamento tra archivi e/o tra risorse.

Il sistema deve inoltre garantire la conservazione a lungo termine del patrimonio digitale.

2 Biblioteche, archivi o collezioni?

Digital library è un termine ambiguo, quasi totalmente assente nei progetti di matrice archivistica o museale: gallerie multimediali, teche, archivi digitali sono i termini alternativi utilizzati.

Numerosi sono i gruppi di ricerca e gli studi che analizzano e definiscono il concetto di *digital library* (Guerrini, Gambari, 2002; Noiret, 2011; Weston 2002; Salarelli,

Tammaro, 2006; Santoro, 2006; ARL² DELOS³, Digital Library Federation⁴; IFLA⁵; IRCDL⁶, TPD⁷).

La maggior parte dei sistemi informativi che forniscono l'accesso a contenuti digitali culturali utilizzano il termine *digital library* nel suo senso più vasto:

Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.⁸

Al di là della terminologia adottata per definire i sistemi informativi di accesso ai patrimoni culturali, è di cruciale importanza restituire le informazioni su provenienza e contesto delle risorse, con la consapevolezza che è necessario fornire un accesso informativo integrato e controllato dagli specialisti relativo ai patrimoni digitali resi disponibili online. Per evitare di cadere nel rumore o nel silenzio informativo in fase di recupero dell'informazione, è necessaria una continua ricerca e sperimentazione per calibrare in sede di acquisizione delle risorse la specificità e il dettaglio scientifico della descrizione con i modelli adottati per la restituzione dei dati, che devono agevolare la facilità d'uso e la comprensione degli strumenti di ricerca.

Questi aspetti sono connessi alla qualità delle rappresentazioni e centrali nella fase di progettazione iniziale. Naturalmente, un certo margine di compromesso è inevitabile, ma esiste un livello minimo accettabile perché non si perda il senso del lavoro descrittivo e si sviluppino soluzioni meno appiattite di quelle offerte dai motori di ricerca oggi disponibili.

Le soluzioni di *digital library* sono valutabili se si circoscrivono i modelli di riferimento e si definiscono categorie specifiche di patrimoni. I primi modelli e le prime definizioni elaborate alcuni anni fa hanno peraltro natura molto generale come risulta anche dal manifesto di DELOS che nel 2007 considerava la *digital library* "as a tool at the centre of intellectual activity having no logical, conceptual, physical, temporal,

² Association of Research Libraries. URL=

<http://www.arl.org/resources/pubs/mmproceedings/126mmappen2.shtml> [ultima visita 18.2.2013].

³ Network of Excellence on Digital Libraries. URL= <http://www.delos.info/>

⁴ Digital Library Federation. URL= <http://diglib.org/> [ultima visita 18.2.2013].

⁵ International Federation of Library Associations and Institutions URL=

<http://archive.ifla.org/VI/5/op/udtop8/udtop8.htm> [ultima visita 18.2.2013].

⁶ Italian Research Conference on Digital Libraries. URL=

<http://ims.dei.unipd.it/websites/ircdl/home.html> [ultima visita 18.2.2013].

⁷ Theory and Practice of Digital Libraries. URL= <http://tpdl.eu/> [ultima visita 18.2.2013]

⁸ Digital Library Federation. URL=

<http://diglib.org/about/dldefinition.htm> [ultima visita 18.2.2013].

or personal borders or barriers on information”, sottolineando la novità di un nuovo paradigma per l’accesso ai patrimoni digitali:

It has moved from a content-centric system that simply organizes and provides access to particular collections of data and information, to a person-centric system that aims to provide interesting, novel, personalized experiences to users. Its main role has moved from static storage and retrieval of information to facilitation of communication, collaboration, and other forms of interaction among scientists, researchers, or the general public on themes that are pertinent to the information stored in the Digital Library. Finally, it has moved from handling mostly centrally-located text to synthesizing distributed multimedia document collections, sensor data, mobile information, and pervasive computing services.⁹

Non sembrano trovare facilmente collocazione e soprattutto univocità di trattamento in questo contesto alcune categorie di materiali informativi, tra cui i patrimoni documentari sedimentati nello svolgimento delle attività di ricerca, in gran parte non organizzati né catalogati, spesso resi disponibili online attraverso processi di digitalizzazione legati ai progetti e alle indagini in corso e difficilmente riconducibili alle categorie conosciute e governate in ambito disciplinare. Se da un lato, in molti casi si tratta di archivi o di raccolte di natura biblioteconomica, in altri, sempre più frequenti, le esigenze di comunicazione online portano i ricercatori e i produttori di risorse a voler e a dover condividere ambiti, tipologie e contenuti diversificati creando le premesse per una significativa contaminazione delle regole descrittive e dei modelli di rappresentazione della conoscenza.

Nel caso dei materiali prodotti in ambito universitario, in particolare, non si tratta quasi mai né di archivi in senso stretto, né di raccolte documentarie coerenti: i materiali spesso sono sottoposti a selezioni finalizzate alla comunicazione, larga parte della documentazione amministrativa è conservata da strutture diverse, a volte dispersa, a volte presente in modo frammentario e comunque raramente oggetto di trattamento descrittivo e ancor meno di pubblicazione; nel caso di ricerche svolte in partecipazione con altri enti una parte della documentazione collegata è conservata in altre strutture e non è accessibile. È impossibile in questi casi ricorrere a standard descrittivi di un solo ambito disciplinare, sia pure quello archivistico che più di altri sembra prestarsi, per l’ampiezza dei termini di riferimento, a comprendere una variegata gamma di produzioni documentarie.

Il termine *collezione* alla fine utilizzato per necessità operative più che per convinzione scientifica nel progetto che qui si presenta è altrettanto ambiguo del concetto di *digital library*, come risulta chiaramente dalla definizione, anzi dalle definizioni proposte dalla Society of American archivists e riprese dalle *Rules for archival description* elaborate dagli archivisti canadesi:

⁹ DELOS, *The Digital Library Manifesto* (2007) URL=
http://www.delos.info/index.php?option=com_odl_open&handle=delos/2006_other_0081&version=1&Itemid=333 [ultima visita 18.2.2013].

“Collection n. ~ 1. A group of materials with some unifying characteristic. - 2. Materials assembled by a person, organization, or repository from a variety of sources; an artificial collection”¹⁰.

Spesso infatti non è una proprietà o caratteristica comune degli oggetti che li tiene insieme, bensì – come indica la seconda definizione del termine – una precisa volontà di assemblare e collezionare documenti, testimonianze e oggetti, altre volte invece una collezione può essere il residuo di un archivio andato disperso o frammentato presso istituzioni diverse.

An artificial accumulation of documents of any provenance brought together on the basis of some commons characteristic, e.g. way of acquisition, subject, language, medium, type of document, name of collector, which may be treated for descriptive purpose as unit under common title (Bureau of Canadian Archivists, 1990, 2008, Appendix D p.3).

Nel caso specifico molti patrimoni attualmente in fase di digitalizzazione e metadattazione per l'esposizione sul portale SDL sono addirittura collezioni nate a posteriori – se così si può dire – da archivi di ricerca, nel senso che si tratta di materiale riordinato e organizzato per la fruizione in un momento successivo alla sua produzione, selezionando una parte della documentazione, che nella maggior parte dei casi non comprende tutti i materiali che sono stati accumulati in fase di espletamento delle proprie attività e funzioni. Molto spesso infatti, come si è già ricordato, il vero e proprio archivio amministrativo e di ricerca di missioni archeologiche, etnologiche o di altro tipo è in carico alle strutture amministrative delle istituzioni e ai ricercatori che hanno partecipato a vario titolo alla Missione. Spesso si tratta di enti, istituzioni, persone di diversa natura, del tutto indipendenti e, in alcuni casi anche appartenenti ad altre nazioni.

Il tema meriterebbe di essere ulteriormente approfondito, così come il contesto attuale di gestione condivisa di patrimoni digitali di diversa provenienza richiederebbe una maggiore attenzione nell'utilizzo di una terminologia appropriata eventualmente anche aggiornata con rigore e creatività. Ma quello che è qui importante sottolineare, relativamente alle esigenze operative che il progetto SDL ha dovuto affrontare, è che tutte le considerazioni fin qui esposte avvalorano la necessità di gestire i processi di digitalizzazione e pubblicazione online dedicando alla qualità delle informazioni di rappresentazione e descrizione molta più attenzione di quanto non si sia fatto nel passato anche in iniziative internazionali di grande rilevanza come Europea. Si tratta in particolare di riconoscere che la contaminazione resa necessaria dai sistemi digitali di comunicazione non può tradursi nell'adozione di un solo modello descrittivo e dei suoi linguaggi di riferimento (Dublin Core? MODS? EAD/EAC?), ma richiede che per ciascuna area informativa e per ciascun dominio si rispetti la specificità funzionale e si adottino le soluzioni che garantiscano la miglior qualità possibile nella restituzione delle informazioni necessarie. La convergenza dei sistemi informativi, per molti versi resa inevitabile dalle strategie di evoluzione del web, non può comportare l'eliminazione delle specificità senza produrre conseguenze negative sulla qualità del trattamento e della restituzione delle fonti. Allo stesso tempo le comunità scientifiche

¹⁰ URL= <http://www2.archivists.org/glossary/terms/c/collection>
[ultima visita 02.03.2013].

e professionali cui compete la responsabilità della comunicazione delle memorie digitali, non possono continuare a dividersi frammentando e indebolendo l'offerta informativa. Gli interventi di mediazione interdisciplinare sono quindi necessari e implicano un notevole grado di complessità, tanto più se non si vuole rinunciare da un lato agli obiettivi della interoperabilità tra sistemi e universi di conoscenza, dall'altro alla specializzazione delle informazioni.

L'esigenza di integrazione tra i diversi domini che si occupano di patrimoni culturali è ormai così sentita nelle comunità di riferimento che nel 2011 le associazioni nazionali operanti nei vari ambiti disciplinari del settore dei patrimoni culturali (AIB¹¹, ANAI¹² e ICOM Italia¹³), hanno dato vita a un coordinamento permanente (MAB¹⁴) per esplorare le prospettive di convergenza tra i mestieri e gli istituti in cui operano i professionisti degli archivi, delle biblioteche, dei musei:

Il Coordinamento si propone come luogo di elaborazione di proposte e di azioni che pongano gli operatori professionali dei beni culturali nella condizione di sviluppare esperienze di collaborazione, di confrontarsi sulle criticità dei singoli ambiti disciplinari e di promuovere soluzioni organizzative, normative, tecnico-scientifiche per gli istituti culturali e per il patrimonio culturale in genere. (Atto costitutivo MAB, 2012, p.1)

Nel caso specifico, con riferimento alle scelte operate nel progetto SDL, ad esempio, allorché si è trattato di definire i modi per documentare le informazioni di contesto delle 'collezioni', si è ritenuto di dover far ricorso agli standard sviluppati in ambito archivistico, sia per quanto riguarda le componenti informative relative al soggetto/ai soggetti che hanno prodotto i materiali e li hanno versati alla SDL, sia e ancor più allorché si tratta di dar conto della specificità di sedimentazione della collezione medesima, delle sue articolazioni, della sua struttura diversificata che prevede partizioni, serie, sezioni diverse. Si è infatti verificato nei fatti che gran parte delle raccolte documentarie prodotte in ambito scientifico richiedono per molti aspetti un trattamento descrittivo di natura archivistica capace di individuare non solo il livello di aggregazione di ciascuna entità, ma anche e soprattutto di definire e restituire in fase di catalogazione e di ricerca le informazioni sul contesto e sulla struttura, foriere di ulteriore arricchimento semantico.

In molti casi le collezioni sono inoltre complessi dinamici, che si incrementano periodicamente di nuove risorse, di inedite relazioni e nuovi vincoli: la struttura della collezione deve quindi rispecchiare la loro complessità originaria e potenziale ed essere aperta a nuove inevitabili acquisizioni nonché a ricongiungersi virtualmente con fondi conservati in altri luoghi ma con cui condivide legami originari.

¹¹ Associazione Italiana Biblioteche URL= <http://aib.it> [ultima visita 02.03.2013].

¹² Associazione Nazionale Archivistica Italiana URL= <http://anai.org/anai-cms/> [ultima visita 02.03.2013].

¹³ International Council of Museum - Comitato Nazionale Italiano URL= <http://icom-italia.org/> [ultima visita 02.03.2013].

¹⁴ Musei Archivi, Biblioteche - Professionisti del patrimonio culturale URL= <http://mab-italia.org/> [ultima visita 02.03.2013].

2.1 Set di metadati, integrazione e interoperabilità

Il set di metadati adottato dal progetto è il MODS¹⁵, verso cui vengono mappati gli standard di descrizione di struttura e contenuto appartenenti a diversi domini (MARC¹⁶, RDA, Dublin Core¹⁷, ISAD, EAD¹⁸, EAC¹⁹, LIDO²⁰)

L'armonizzazione delle pratiche catalografiche viene realizzata mediante un'attività continua di studio e di confronto tra diverse tradizioni di dominio, e con specifici accordi con le altre istituzioni nazionali deputate alla conservazione e alla fruizione del patrimonio culturale: ICAR²¹, ICCU²², ICCD²³. Come già in parte sottolineato, infatti, la normalizzazione non deve essere confusa con l'omologazione, né l'integrazione con la convergenza. Va intesa piuttosto – almeno questa è stata la scelta operata nel progetto SDL – come la ricerca di un minimo comune denominatore che permetta l'interrogazione cumulativa di materiali eterogenei descritti in base ai differenti domini disciplinari, senza per questo impoverire le informazioni, che sono poi restituite per intero nell'esplorazione della risorsa.

Non sono naturalmente poche le aree che hanno richiesto e che ancora richiedono una riflessione multidisciplinare finalizzata ad adottare soluzioni convincenti e a risolvere ambiguità interpretative, a cominciare dalla chiara definizione di ciò che una *digital library* descrive, con riferimento ad esempio alla distinzione tra l'oggetto originario e la *copia digitale* realizzata attraverso il processo di digitalizzazione e destinata alla pubblicazione online.

2.2 Oggetto fisico e oggetto digitale

Gli aggregatori di contenuti digitali che non siano solo cataloghi on line o inventari di patrimoni posseduti hanno la necessità di distinguere l'oggetto fisico dall'oggetto digitale e indicare quali metadati si riferiscono alla risorsa originaria e quali alla sua rappresentazione digitale.

¹⁵ Metadata Object Description Schema URL= <http://loc.gov/standards/mods/> [ultima visita 18.2.2013].

¹⁶ Machine-Readable Cataloging URL= <http://loc.gov/marc/> [ultima visita 18.2.2013].

¹⁷ URL= <http://dublincore.org/> [ultima visita 18.2.2013].

¹⁸ Encoded Archival Description URL= <http://loc.gov/ead/> [ultima visita 18.2.2013].

¹⁹ Encoded Archival Context URL=
<http://www2.archivists.org/standards/encoded-archival-context-corporate-bodies-persons-and-families-eac-cpf>
URL= <http://eac.staatsbibliothek-berlin.de/> [ultima visita 18.2.2013].

²⁰ URL= <http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/> [ultima visita 18.2.2013].

²¹ Istituto Centrale per gli Archivi URL= <http://icar.beniculturali.it> [ultima visita 18.2.2013].

²² Istituto centrale per il catalogo unico URL= <http://iccu.sbn.it/> [ultima visita 18.2.2013].

²³ Istituto centrale per il catalogo e la documentazione URL=
<http://iccd.beniculturali.it/> [ultima visita 18.2.2013].

Il rigore scientifico riguardo l'oggetto della descrizione e il mantenimento della catena di conservazione possono realizzarsi solo dando conto in maniera puntuale dei metadati che si riferiscono alla risorsa analogica e quelli riferiti al surrogato digitale: il ProvidedCHO Provided Cultural Heritage Object rappresentato dalla Web Resource dell'EDM (Europeana Data Model)²⁴ adottato dal portale Europeana, l'oggetto fisico reale e la sua rappresentazione digitale.

Il portale Europeana²⁵, in quanto aggregatore di metadati, una sorta di catalogo di secondo livello, raccoglie i dati sui patrimoni di diverse istituzioni culturali sul territorio europeo, senza essere responsabile né delle risorse originarie né della conservazione a lungo termine dei materiali digitali. Ha potuto quindi sciogliere il nodo in questione con una certa facilità, anche se è giunto a definire con chiarezza il problema solo recentemente.

Nel caso invece di una *digital library* che sia aggregatore ma anche produttore e collettore di archivi, collezioni e materiali – digitalizzati o nativi digitali – che forma nell'esercizio della sua attività di ricerca e di cui ha quindi la responsabilità scientifica e di conservazione digitale, i nodi teorici da affrontare sono ancor più complessi e non sempre facilmente distinguibili. La doppia funzione di ente produttore e depositario della catena di conservazione degli originali e di canale di comunicazione di patrimoni non propri e di cui non ha la responsabilità della produzione, peraltro posseduti se non in digitale, costituisce infatti un fattore di complessità a vari livelli: concettuali al momento della catalogazione, ma anche organizzativi in sede di definizione dei flussi di lavorazione e di responsabilità.

Nel nostro caso, SDL non è l'istituzione deputata alla conservazione e all'accesso alle collezioni fisiche: è un catalogo per così dire di 'secondo livello' e può rimandare al catalogo dell'istituzione custode dei relativi oggetti analogici o a un'interfaccia di interrogazione specifica per database complessi. In questo caso, peraltro, la catena di conservazione va preservata, e anche se l'accesso ai documenti originari non è regolato da SDL, è tuttavia indispensabile stabilire e rendere visibile il raccordo con il catalogo 'primario'. Tale raccordo avviene a più livelli: innanzitutto tramite la collocazione del documento fisico, il cui inserimento e aggiornamento va curato dalle strutture che forniscono/versano i materiali cui compete garantirne la disponibilità e l'accesso. In secondo luogo, tale raccordo si esprime mediante la descrizione, secondo standard archivistici desunti dal modello adottato dal Sistema archivistico nazionale e coerente con gli standard EAD/EAC, del soggetto produttore e del soggetto versante come si dirà in seguito.

Questa distinzione dei diversi piani concettuali permette inoltre di preservare le informazioni sulla provenienza e di aggregare materiali diversi dando conto nei metadati descrittivi e gestionali-amministrativi se i dati stessi si riferiscono all'oggetto fisico, alla sua rappresentazione digitale, o se invece le risorse sono native digitali.

Quindi i possibili oggetti della catalogazione sono:

- oggetti *born digital*

²⁴ URL= <http://europeana.eu/schemas/edm/> URL=
<http://europeanconnect.eu/results-and-resources.php?page=1>
[ultima visita 18.2.2013].

²⁵ URL= <http://europeana.eu/portal/> [ultima visita 18.2.2013].

- rappresentazioni digitali di originali analogici, con link alla localizzazione della risorsa originaria.
- riproduzioni di oggetti che non possiedono un proprio catalogo o inventario o che non siano conservati in uno specifico museo o istituzione (ad esempio la riproduzione fotografica di un paesaggio o di uno scavo), che quindi non avranno il collegamento alla catalogazione dell'esemplare fisico originario.

In questa logica vengono sciolte le ambiguità riscontrate nei campi portatori di maggiore sovrapposizione semantica: autore o *creator*, data e luogo. Queste criticità sono particolarmente rilevanti nel caso di oggetti fisici di natura non documentaria ma di diverso tipo: reperti archeologici, opere d'arte, oggetti museali.

Il concetto di autore è inevitabilmente ambiguo nel contesto multidisciplinare in cui si opera, non solo per quanto concerne l'individuazione dell'agente (autore della rappresentazione fotografica o dell'oggetto rappresentato, ambito quest'ultimo che ricade senz'altro nella semantica, ma che spesso viene imprecisamente metadato come *creator*) ma anche nel caso di patrimoni di tipo archivistico o museale in cui può frequentemente non coincidere con l'autore dell'oggetto descritto, trattandosi piuttosto di autore nell'accezione di soggetto produttore dell'archivio o dell'entità che ha raccolto e organizzato una collezione di oggetti o rappresentazioni documentarie (analogiche o digitali) di attività di ricerca.

L'assunto di partenza è che va distinto nettamente l'ambito della risorsa originaria analogica (laddove esistente) e le relative competenze di conservazione e accesso ai documenti che sono in capo alle strutture detentrici; e il piano dell'oggetto digitale, che costituisce un livello di descrizione autonomo, esclusivamente semantico, da collegare all'eventuale descrizione dell'originale. Tale distinzione deve peraltro essere innanzitutto definita nella fase preliminare di descrizione della collezione e delle sue partizioni. È infatti in quella sede che è possibile chiarire con maggiore coerenza (anche con il supporto del responsabile scientifico della collezione) obiettivi e natura del progetto di digitalizzazione e stabilire ruoli e qualifiche degli attori che a vario titolo sono presenti nella produzione delle singole risorse, non diversamente del resto da quanto avviene nella descrizione di materiali d'archivio.

La costruzione di un solido impianto concettuale intorno all'oggetto della descrizione è del resto alla base dell'architettura dei contenuti e permette di creare un catalogo che in fase di ricerca restituisca risultati pertinenti e riduca il rumore informativo al minimo.

2.3 Descrizione a più livelli dei complessi documentari

Nel sistema progettato per la SDL, la descrizione delle collezioni – a differenza di quanto previsto nella gran parte delle *digital library* – avviene rispettando e rappresentando adeguatamente la struttura a più livelli del patrimonio documentario e le serie archivistiche, in armonia con gli standard di dominio e in considerazione del fatto che il concetto di collezione qui utilizzato è in molti casi, come si è già sottolineato, riconducibile a quello di archivio, anche se le aggregazioni archivistiche descritte non sono sempre quelle originarie.

È del resto necessario preservare il concetto di descrizione a più livelli per rappresentare patrimoni informativi/documentari portatori di strutture interne profonde e complesse e relazioni esterne molteplici.

Nello specifico il sistema di *digital library* gestisce con cura il sistema di relazioni che collegano le partizioni della collezione e i rapporti tra le collezioni e i diversi soggetti che ne sono all'origine o che si sono avvicinati nella loro gestione e custodia, riunendo concettualmente fondi che sono fisicamente dispersi (Yeo, 2012).

2.4 Descrizione dei soggetti produttori e versanti e documentazione di progetto

Se, come abbiamo visto, nel concetto di collezione è intrinseca l'intenzionalità nel riunire i materiali e sono necessarie delle scelte per selezionare e organizzare logicamente le risorse, in un progetto di digitalizzazione è di fondamentale importanza dare conto della storia e della provenienza della collezione e dichiarare i tipi di intervento e i criteri che li hanno guidati. Solo in questo modo è possibile rendere trasparenti questi processi e intellegibili le fonti documentarie e la storia culturale di cui sono portatori.

Gli apparati informativi di corredo si rendono indispensabili, oltre che per garantire ricchezza informativa e contestualizzazione dei patrimoni, anche per la certificazione della provenienza del documento e delle responsabilità in capo a singoli individui o istituzioni. Infatti "l'autenticità del documento può essere dimostrata o presunta se le informazioni di contesto amministrativo e documentario si sono mantenute e sono ricostruibili e ricercabili" [Guercio, 2010, p.28]. Questo è ancor più vero in ambiente digitale, dove è opportuno esplicitare ogni tipo di informazione fornendo una descrizione accurata della provenienza e della storia delle collezioni alle quali appartengono i singoli oggetti. La documentazione di contesto e di progetto è essenziale per garantire autorevolezza scientifica alla catena di conservazione e integrità e autenticità alle fonti documentarie: ambedue elementi necessari per stabilire un rapporto fiduciario tra i *service provider* e gli utenti.

In virtù di queste considerazioni particolare attenzione è stata posta nell'assicurare la descrizione dei soggetti produttori e di quelli versanti, che nel contesto della SDL acquistano peso giuridico oltre che un ruolo organizzativo. Nello specifico il sistema di *digital library* gestisce con cura il sistema di relazioni che collegano le partizioni della collezione e i rapporti tra le collezioni e i diversi soggetti che ne sono all'origine o che si sono avvicinati nella loro gestione e custodia.

Nel sistema SDL i progetti di digitalizzazione sono stilati in collaborazione tra gli esperti di dominio archivistico e bibliotecario e i responsabili scientifici e tecnici delle collezioni. Oltre a una descrizione dell'archivio/collezione effettuata secondo le norme del SAN²⁶ e nel rispetto degli standard ISAD(G)²⁷/ISAAR(CPF)²⁸, ogni comples-

²⁶ Sistema Archivistico Nazionale URL= <http://san.beniculturali.it> [ultima visita 18.2.2013].

²⁷ General International Standard Archival Description URL= <http://ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html> [ultima visita 18.2.2013].

²⁸ International Standard Archival Authority Record for Corporate Bodies, Persons and Families URL= <http://ica.org/10203/standards/isaar-cpf-international-standard-archival-authority-record-for->

so documentario sottoposto a digitalizzazione, metadattazione e/o normalizzazione dei dati, è corredato da una scheda di progetto che fornisce informazioni sulla storia dell'archivio/collezione, sulle responsabilità tecniche, scientifiche ed editoriali coinvolte, sul soggetto produttore e sul soggetto versante, sulla consistenza della raccolta, sui criteri di selezione per la digitalizzazione e sulle fasi e gli incrementi previsti. Il progetto scientifico di digitalizzazione o di recupero di patrimoni digitali costituisce parte integrante della descrizione, una sorta di nota archivistica in grado di fornire elementi cruciali sulla storia del contesto tecnologico, istituzionale, amministrativo e sulla provenienza intesa anche come definizione di responsabilità specifiche per la produzione e gestione di patrimoni documentari digitali.

2.5 Normalizzazione del linguaggio: authority file e thesauri

Al fine di garantire interoperabilità e coerenza nella descrizione, la definizione e la gestione dei metadati nel progetto SDL ha implicato una notevole attenzione in termini di qualità dei processi di normalizzazione. Si è fatto perciò costante ricorso a authority file e thesauri la cui tipologia è stata individuata con riferimento agli standard internazionali, coerentemente con i diversi materiali descritti e nell'ottica dei linked data: TGN²⁹ e Geonames³⁰ per i luoghi, VIAF³¹ per i nomi di persona ed ente, PICO 4.3³² e Marc Code List for Relators³³ per i ruoli, PICO e il Nuovo soggetto di Firenze³⁴ per la soggettazione, con la possibile integrazione di vocabolari controllati disciplinari.

Grazie all'inserimento controllato dei nomi, dei descrittori, delle date e dei luoghi è possibile perciò creare dei punti di accesso realmente efficaci nell'indicizzazione e nella contestualizzazione spazio-temporale dei contenuti.

2.6 Catalogazione descrittiva, semantica e relazionale

La letteratura di settore, la ricerca e l'applicazione sperimentale di piattaforme, modelli di dati e ontologie per la realizzazione di sistemi integrati di accesso al patrimonio culturale e della ricerca stanno attraversando una fase di riflessione, grande fer-

corporate-bodies-persons-and-families-2nd-edition.html [ultima visita 18.2.2013].

²⁹ Getty Thesaurus of Geographic Names URL=
<http://getty.edu/research/tools/vocabularies/tgn/> [ultima visita 18.2.2013].

³⁰ Geonames URL= <http://geonames.org/> [ultima visita 18.2.2013].

³¹ Virtual International Authority File URL= <http://viaf.org/> [ultima visita 18.2.2013].

³² Portale Italiano della Cultura on line URL=
http://culturaitalia.it/opencms/export/sites/culturaitalia/attachments/thesaurus/4.3/thesaurus_4.3.0.skos.xml [ultima visita 18.2.2013].

³³ URL= <http://loc.gov/marc/relators/relaterm.html> [ultima visita 18.2.2013].

³⁴ URL= <http://thes.bncf.firenze.sbn.it/> [ultima visita 18.2.2013].

mento e transizione. Negli ultimi quindici anni sono stati realizzati numerosi progetti di digitalizzazione e/o aggregazione di archivi a ogni livello: locale, nazionale ed europeo, come nel caso di Europeana. Tuttavia, non si è stati ancora capaci di sfruttare appieno le potenzialità degli strumenti digitali. Nella gran parte dei progetti di *digital library* finora realizzati si è ad esempio rinunciato a mantenere e integrare le soluzioni descrittive delle diverse tradizioni disciplinari consolidate, privilegiando scelte 'parziali' che hanno in molti casi reso disponibili collezioni digitali dotate di limitate o del tutto inesistenti informazioni di contesto, senza indicazioni sulle aggregazioni e sulle relazioni interne, con descrizioni limitate alla sola collezione e ai singoli oggetti. Un buon catalogo di biblioteca digitale dovrebbe invece esplicitare anche il tessuto di relazioni implicite fra i documenti che descrive ed espone, superando i limiti del pensiero lineare gutenberghiano e ri-pensando il catalogo come una rete di relazioni semantiche e spazio-temporali.

La Library of Congress ha avviato una profonda riflessione sull'opportunità di superare il concetto ormai riduttivo di descrizione elaborato in ambito bibliotecario e orientato all'esemplare e al formato fisico, per orientarsi verso una descrizione di tipo semantico e relazionale che tenga conto di diverse tipologie di materiale, superando i limiti che impone il formato Marc, nato quaranta anni fa in ambito bibliotecario e più volte rielaborato, ma ormai insufficiente per rappresentare materiali non bibliografici e relazioni complesse, come si legge nel report *On the record* della Library of Congress:

Different communities of bibliographic practice have grown up around different resource types: library collections of books and journals; archives; journal articles; and museum objects and images. As these resources and others become increasingly accessible through the Web, separation of the communities of practice that manage them is no longer desirable, sustainable, or functional. Bibliographic control is increasingly a matter of managing relationships — among works, names, concepts, and object descriptions — across communities. Consistency of description within any single environment, such as the library catalog, is becoming less significant than the ability to make connections between environments [...].

(Library of Congress, 2008, p. 9)

La rivoluzione digitale ci permette e ci impone di superare una catalogazione che sia attenta alla semantica in maniera disomogenea (sia per adozione di strumenti di controllo terminologico che per applicazioni pratiche) per portarci verso una catalogazione focalizzata sulla semantica e sulle relazioni di contesto, semantiche e strutturali e sia in grado di normalizzare i vocabolari integrandoli al fine di assicurarne anche la specializzazione.

Lo scopo è ricondurre la conoscenza a uno stadio di ricchezza che possa servire a descriverne le varie forme e manifestazioni in modo univoco, consentendo la conoscenza di materiali eterogenei senza perderne le informazioni di provenienza ed esplicitando le relazioni implicite tra le risorse.

La sfida che SDL ha voluto raccogliere risponde all'esigenza, da tempo dichiarata anche se scarsamente perseguita nei diversi progetti di biblioteche digitali, di assicurare un accesso alle risorse arricchito di informazioni descrittive grazie a nuove forme

di intermediazione scientifica e fornite dagli specialisti di dominio, da coloro cioè che sono chiamati a "istituire, coordinare e connettere sintassi di segni" (Serrai, 1997, p. 29, nota 1).

3 Conclusioni

I sistemi informativi digitali permettono organizzare i materiali logicamente e semanticamente, restituendo i legami di contesto e la cornice di significato delle risorse. La normalizzazione della terminologia dei descrittori, l'utilizzo dei linguaggi e delle grammatiche del web semantico e la modellazione della realtà e degli oggetti per mezzo delle ontologie e degli standard di descrizione, possono coadiuvare l'ambizione che in fase di ricerca venga restituita un'informazione di qualità, controllata e non ridondante.

Il materiale grezzo acquista significato solo se inserito in una rete semantica e di riferimenti spazio-temporali, che a sua volta viene arricchita dalle singole risorse.

Una *digital library* non è semplicemente la riproposizione in forma elettronica, organizzata linearmente, del catalogo cartaceo, né un semplice deposito di risorse digitali, bensì è un complesso strumento di organizzazione della conoscenza. È necessario superare il concetto di catalogo elettronico come semplice trasposizione in digitale di collezioni fisiche, e sperimentare linguaggi, strumenti e formalismi atti a descrivere ed esplorare l'universo rappresentato fornendo all'utente, oltre alle informazioni sulle singole risorse e sulla loro semantica, le informazioni di contesto sull'ambiente di produzione collocato nelle coordinate spazio-temporali che gli sono proprie. È a questo complesso insieme di obiettivi che il progetto SDL si è ispirato e ha l'ambizione di contribuire.

4 Bibliografia

- Arms William Y (2012) *The 1990s: The Formative Years of Digital Libraries* in «Library Hi Tech», vol 30, no. 4 URL=<http://www.emeraldinsight.com/journals.htm?articleid=17065297> [ultima visita 02.03.2013].
- Bureau of Canadian Archivists (1990, revised 2008) *Rules for Archival Description, Appendix D, Glossary* URL=http://www.cdncouncilarchives.ca/RAD/RADComplete_July2008.pdf [ultima visita 02.03.2013].
- Calvo M., Ciotti F., Roncaglia G., Zela Marco A. (2004). *Internet 2004*, Laterza. URL=<http://www.marcocalvo.it/libri/internet/home/index.htm> [ultima visita 02.03.2013].
- Capaldi D., Ilardi E., Ragone G. (2011). *I cantieri della memoria*, Liguori.
- Curral J., Moss M., Stuart S. (2004) *What is a collection?* «Archivaria», no 58 URL=<http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12480/13594> [ultima visita 02.03.2013].
- DELOS, *The Digital Library Manifesto* (2007) URL=http://www.delos.info/index.php?option=com_odl_open&handl

- e=delos/2006_other_0081&version=1&Itemid=333 [ultima visita 02.03.2013].
- Library of Congress, *Bibliographic Framework Transition Initiative*. URL= <http://loc.gov/marc/transition> [ultima visita 02.03.2013].
- Library of Congress (2008). *On the record report of the Library of Congress Working Group on the Future of Bibliographic Control*. URL= <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf> [ultima visita 02.03.2013].
- Yeo G. (2012). *The Conceptual Fonds and the Physical Collection*. «Archivaria», no. 73 URL= <http://journals.sfu.ca/archivar/index.php/archivaria/article/view/13384/14691> [ultima visita 02.03.2013].
- Guercio M. (2010). *Archivistica informatica*, Carocci.
- Guerrini M., Gambari S. (2002). *Definire e catalogare le risorse elettroniche* Editrice Bibliografica. URL= <http://eprints.unifi.it/archive/00001949/> [ultima visita 02.03.2013].
- MAB (2012). *Atto costitutivo* URL= http://www.mab-italia.org/images/pdf/MAB-Naz_Atto-costitutivo_2012-06-12.pdf
- Metiteri F., Ridi R. (2007). *Biblioteche in rete*, Laterza. URL= <http://laterza.it/bibliotecheinrete/> [ultima visita 02.03.2013].
- Noiret S. (2011). *Storia digitale* in Perilli L., Fiormonte D., a c. di, *La macchina nel tempo*, Le Lettere, pp. 201-258.
- Ragone (2012). *Introduzione ai lavori*. Quaderni DigiLab», no 2 URL= http://digilab-epub.uniroma1.it/index.php/Quaderni_DigiLab/article/view/17/15 [ultima visita 02.03.2013].
- Salarelli A., Tamaro A. (2006). *La biblioteca digitale*, Editrice Bibliografica.
- Santoro M. (2006). *Biblioteche e innovazione*, Editrice Bibliografica.
- Serrai A. (1997). *Introduzione*. In: Storia della bibliografia, Bulzoni, vol.7.
- Weston P.G. (2001). *Catalogazione bibliografica: dal formato MARC a FRBR*. «Bollettino AIB», vol 41, no 3 URL= <http://bollettino.aib.it/article/view/8536/7616#34> [ultima visita 02.03.2013].
- Weston P.G. (2002). *Il catalogo elettronico*, Carocci.
- Vitali S. (2004). *Passato digitale*, Bruno Mondadori.