

Maria Teresa Biagetti

Le biblioteche digitali

Tipologie, funzionalità
e modelli di sviluppo

FrancoAngeli

Informazioni per il lettore

Questo file PDF è una versione gratuita di sole 20 pagine ed è leggibile con



La versione completa dell'e-book (a pagamento) è leggibile con Adobe Digital Editions. Per tutte le informazioni sulle condizioni dei nostri e-book (con quali dispositivi leggerli e quali funzioni sono consentite) consulta [cliccando qui](#) le nostre F.A.Q.



I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio “Informatemi” per ricevere via e.mail le segnalazioni delle novità.

Maria Teresa Biagetti

Le biblioteche digitali

Tipologie, funzionalità
e modelli di sviluppo

con scritti di

Roberto Raieli - Antonella Iacono -
Antonella Trombone - Simona Turbanti

FrancoAngeli

Il volume è stato pubblicato grazie al contributo del Progetto di Ateneo “Evoluzione delle biblioteche digitali: funzionalità della ricerca e tecnologie semantiche”, finanziato da Sapienza Università di Roma.

Copyright © 2019 by FrancoAngeli s.r.l., Milano, Italy.

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito www.francoangeli.it.

Indice

Premessa	pag. 7
Introduzione	» 9
Parte prima Tipologie, modelli e funzionalità per la ricerca	
1. Le biblioteche digitali: uno scenario composito	» 21
1. I repositories istituzionali e disciplinari	» 23
2. I depositi di riviste	» 28
3. Web archiving	» 30
2. Modelli tradizionali e innovativi	» 43
1. Le biblioteche digitali come repositories di documenti	» 44
1.1 American memory	» 44
1.2 Gallica	» 46
1.3 La Biblioteca digitale italiana	» 50
1.4 National digital library of China	» 53
2. Portali di accesso alle raccolte digitali	» 55
2.1 The Europeana collections	» 55
2.2 The World digital library	» 59
2.3 The Digital public library of America	» 61
3. Progetti di digitalizzazione su vasta scala	» 63
3.1 Google Books	» 63
3.2 HathiTrust	» 64
3.3 Open content alliance	» 66
3.4 I progetti di digitalizzazione di Internet Archive	» 66
3.5 The Universal digital library	» 68
4. Biblioteche digitali per l'analisi testuale: Perseus digital library	» 68
4.1 Funzionalità per la ricerca	» 71
5. Aspetti legali e diritto d'autore per le biblioteche digitali	» 73
5.1 Le opere "orfane"	» 78
5.1.1 L'infrastruttura Arrow	» 80

3. Funzionalità per la ricerca nelle biblioteche digitali	pag. 83
1. La ricerca collaborativa	» 94
4. Gli sviluppi delle biblioteche digitali	» 95
1. Verso la personalizzazione dei servizi	» 97
2. Un nuovo modello: the Delos Project	» 100
3. Il progetto Sekt	» 108
4. Le ontologie	» 110
4.1 Le ontologie bibliografiche	» 114
5. Le biblioteche digitali arricchite con funzionalità semantiche e “social”	» 119
1. Un modello avanzato di repository: Fedora	» 124
2. I progetti Bricks e Simile	» 126
3. Jerome Digital Library	» 128
3.1 L’ontologia MarcOnt	» 132
4. Callimachus Digital Library	» 133
5. Le biblioteche digitali semantiche con funzioni “social”	» 134
5.1 Le annotazioni	» 134
5.2 Supporto automatico per le annotazioni: il prototipo Pirates	» 137
5.3 Servizi “social” in Jerome Digital Library	» 142
5.3.1 Le funzionalità per la ricerca	» 144
6. Spazi informativi per il patrimonio culturale	» 145
1. Interoperabilità tra fonti diverse	» 145
1.1 L’ontologia Cidoc-crm	» 146
2. Il portale CultureSampo	» 153
Considerazioni conclusive	» 155
Riferimenti bibliografici	» 161
Parte seconda	
Funzionalità avanzate di ricerca e sviluppi “social”	
1. Potenzialità della ricerca multimediale nelle biblioteche digitali , di <i>Roberto Raieli</i>	» 177
2. Le competenze dell’utente nello sviluppo delle digital libraries tramite il Crowdsourcing , di <i>Antonella Iacono</i>	» 207
3. Sistemi di raccomandazione e filtraggio collaborativo nella ricerca bibliografica , di <i>Antonella Trombone e Simona Turbanti</i>	» 229
Indice tematico	» 251

Premessa

Il volume raccoglie i risultati delle ricerche svolte nell'ambito del progetto finanziato da Sapienza Università di Roma nel triennio 2015-2018 "Evoluzione delle biblioteche digitali: funzionalità della ricerca e tecnologie semantiche", di cui sono stata responsabile scientifico.

La prospettiva che ho adottato per esaminare le biblioteche digitali è ampia e permette di includere all'interno di questa categoria non soltanto i depositi di oggetti digitalizzati dagli istituti della memoria, ma anche i repositories istituzionali per i lavori prodotti dai ricercatori e l'attività di *web archiving*, che in tutto il mondo negli ultimi anni ha ricevuto un significativo sviluppo. La prima fase della ricerca è stata quindi dedicata ad ampliare le tipologie di raccolte digitali da sottoporre ad analisi, dando spazio a quelle che risultano ancora oggi meno indagate. È stato inoltre necessario analizzare le caratteristiche delle più significative biblioteche digitali di impianto tradizionale e innovativo realizzate dalle istituzioni della memoria, senza però includere le raccolte archivistiche, la cui analisi richiede competenze specifiche, approfondire gli aspetti legali della digitalizzazione e mettere in evidenza le funzionalità offerte alla ricerca nei repositories di oggetti digitali.

Gli obiettivi specifici del progetto volto ad analizzare le caratteristiche e gli sviluppi delle biblioteche digitali prevedevano di saggiare le funzionalità di tipo semantico applicate alle biblioteche digitali formate da raccolte testuali e in particolare l'uso delle ontologie per l'indicizzazione e per il trattamento delle query, e di analizzare la realizzazione di portali culturali come spazi in cui vengono integrati i contenuti eterogenei del patrimonio culturale. Si sono privilegiate quindi le analisi dei nuovi modelli e dei prototipi di biblioteca digitale, e si è cercato di far emergere le funzioni innovative che in alcuni casi sono permesse, come quelle che consentono le annotazioni da parte degli utilizzatori.

Nell'ambito del progetto sono stati inoltre approfonditi alcuni aspetti significativi che riguardano l'universo delle biblioteche digitali, e che costituiscono oggi campi di ricerca di grande interesse e in continua evoluzione. Sono stati infatti indagati i problemi legati alla ricerca degli oggetti multimediali, delle risorse visive e audiovisive, che costituiscono gran parte delle collezioni digitali, e i possibili sviluppi dovuti all'applicazione del Multimedia Information Retrieval. Sono state analizzate le potenzialità offerte dalla creazione di contenuti delle digital libraries anche su base collaborativa, attraverso l'attività di *crowdsourcing*, impiegata nella correzione dei testi digitalizzati con l'uso di software Ocr, nella creazione di link, nella categorizzazione e nell'attività di tagging. Un ulteriore approfondimento è stato riservato ad uno degli aspetti che hanno caratterizzato il processo di progressiva personalizzazione dei servizi offerti dalle digital libraries, e cioè i sistemi di *recommendation*, dei quali viene presentata una rassegna significativa, basati sui profili degli utenti, sui contenuti dei documenti e sulle valutazioni espresse dai lettori.

L'ambito di ricerca dedicato alle biblioteche digitali, nelle diverse forme in cui esse si presentano, è in continua evoluzione e questi scritti sono il risultato di un primo esame di alcuni aspetti e di alcune problematiche che le riguardano, in particolare di quelle che derivano dall'applicazione delle tecnologie del Semantic web, e sicuramente non esauriscono il complesso delle analisi possibili.

Introduzione

L'universo delle biblioteche digitali comprende diverse tipologie, dai repositories nazionali di documenti digitalizzati, fruibili liberamente, ai depositi istituzionali predisposti e gestiti dagli Atenei per accogliere la produzione scientifica dei ricercatori, alle raccolte librerie digitalizzate dalle singole biblioteche e dagli istituti archivistici e, secondo l'opinione di molti studiosi del settore, annovera al suo interno anche le piattaforme per l'e-learning su cui i docenti pubblicano testi, documenti, monografie, per agevolare l'apprendimento a distanza.

Se consideriamo esclusivamente le raccolte librerie digitalizzate, si può affermare che questo settore di attività trovi la sua collocazione all'interno di quel vasto spazio scientifico di ricerca e di applicazioni che costituisce l'elemento di connessione tra la Biblioteconomia e la Scienza dell'informazione, o Library and Information Science, e che si serve del fondamentale apporto tecnologico della Computer Science.

Il legame con la Biblioteconomia è evidente quando si consideri l'importanza sociale che i diversi progetti di digitalizzazione del patrimonio librario rivestono e quando si rifletta sul fatto che gli istituti bibliotecari che conservano il patrimonio sono coinvolti nella scelta delle raccolte da digitalizzare. Le biblioteche digitali permettono di realizzare raccolte selezionate, che rispondono a precisi requisiti appropriati alle necessità espresse dalle diverse categorie di utilizzatori.

La forte connessione con la Scienza dell'informazione riguarda invece in particolare le modalità di accesso intellettuale alla conoscenza registrata nei documenti, l'organizzazione dei contenuti informativi, l'interazione degli utenti con le interfacce dei sistemi di ricerca e, inoltre, la fruizione dei contenuti informativi multimediali.

Il campo scientifico che studia le problematiche relative alle biblioteche digitali presenta una natura interdisciplinare: gli specialisti di Computer Science si occupano degli aspetti tecnologici e dello sviluppo degli algo-

ritmi, gli studiosi del settore della Library and Information Science si occupano delle questioni relative ai contenuti, all'organizzazione delle funzionalità della ricerca e all'analisi dei comportamenti degli utenti, mentre competono ai sociologi lo studio dei contesti sociali all'interno dei quali avviene la fruizione.

La visione che i bibliotecari hanno della biblioteca è particolarmente ampia e riguarda le diverse funzioni relative alla scelta delle pubblicazioni da offrire agli utenti, all'organizzazione dei documenti, alla loro conservazione e alle modalità che ne permettono l'accesso. Con la nascita delle biblioteche digitali i bibliotecari hanno praticamente ampliato la loro concezione originaria della biblioteca fino ad includere il nuovo ambiente digitale, e in questa visione sono impliciti i significati di conservazione nel tempo delle risorse informative, in continuità con la visione della biblioteca tradizionale. In questa prospettiva l'accento viene posto sulla costituzione di una raccolta di documenti congruenti rispetto ad una utenza definita, viene dato rilievo alla funzione sociale della conoscenza, e la biblioteca digitale è concepita sostanzialmente come una estensione della biblioteca tradizionale. I computer scientists, invece, manifestano una visione più ristretta di biblioteca, legata al concetto di database e all'Information Retrieval, quindi sono particolarmente interessati all'organizzazione dei dati e alle funzionalità dedicate a fornire l'accesso alle risorse¹.

Le biblioteche digitali sono sistemi complessi e la loro realizzazione coinvolge un numero rilevante di componenti, in primo luogo le raccolte di testi e di immagini e, di eguale importanza, l'infrastruttura tecnica per il sistema informativo che si decide di realizzare. Il progetto culturale su cui l'impresa si basa riveste tuttavia un ruolo centrale e fornisce l'impronta scientifica della realizzazione. A differenza delle biblioteche reali, che possono essere state istituite nei secoli passati e presentare raccolte che si sono stratificate nel tempo, non sempre sviluppatasi in modo armonico, non sempre del tutto congruenti con le necessità di una comunità, una raccolta di oggetti digitali può essere costruita tenendo conto precisamente delle esigenze espresse dalle necessità dell'utenza e le opere potranno quindi essere selezionate in base alle necessità dei fruitori cui la biblioteca digitale si rivolge, e agli obiettivi scientifici che essa si prefigge.

Come ha messo in evidenza Christine Borgman (2000), che ha preso parte a molte attività che hanno riguardato le biblioteche digitali degli Stati Uniti, le biblioteche digitali giocano un ruolo fondamentale nella comunicazione accademica, hanno assunto un significativo ruolo sociale e sono

¹ Christine L. Borgman, oggi professore di Information Studies all'Università della California, Los Angeles (Ucla), ha sostenuto questa posizione in diversi suoi scritti, in particolare in (1999), disponibile a: <https://escholarship.org/uc/item/7c55m1xf>.

entrate a far parte del circuito della comunicazione scientifica che, oltre a riguardare gli studiosi e le istituzioni bibliotecarie, interessa anche il mondo dell'editoria. Nell'era digitale la comunicazione scientifica ha affrontato rilevanti cambiamenti, che hanno riguardato non solo le modalità di fare ricerca da parte del singolo ricercatore, con l'incremento della possibilità di scambiare riflessioni e informazioni, ma anche la diffusione della possibilità di pubblicare in modalità open access.

Nell'ambito delle ricerche e dei gruppi di lavoro finanziati a partire dal 1991 dalla US National Science Foundation (Nsf)², hanno rivestito particolare importanza quelli avviati nel 1994 e dedicati alla Digital Libraries Initiative, un ampio progetto sostenuto dalla stessa National Science Foundation, da Darpa³ e Nasa⁴ con l'obiettivo di elaborare strumenti per raccogliere, immagazzinare, organizzare e rendere fruibili risorse informative in formato digitale⁵.

Tra il 15 e il 17 febbraio 1996 si tenne un importante workshop nell'ambito della Digital Libraries Initiative presso la University of California di Los Angeles (Ucla). Le discussioni ruotarono principalmente attorno al riconoscimento del fatto che la realizzazione delle biblioteche digitali comporti rilevanti problemi di ordine sociale, la cui risoluzione richiede non solo risorse tecnologiche ma soprattutto specifiche capacità e competenze specialistiche di ambito biblioteconomico.

Il rapporto finale del workshop, firmato da Christine Borgman (1996)⁶, mette in evidenza in modo particolarmente efficace la dialettica creatasi all'interno dei partecipanti al gruppo di ricerca. Vengono offerte, infatti, due definizioni, e due visioni, complementari di biblioteca digitale: una si focalizza sulla realizzazione tecnica delle risorse elettroniche, sulle funzionalità per la ricerca e sulla possibilità di utilizzare le informazioni, nell'ottica del miglioramento dei sistemi di "storage and information retrieval" che trattano i dati digitali, testi, immagini statiche e dinamiche, suoni,

² National Science Foundation è l'agenzia governativa statunitense, istituita dal Congresso nel 1950, con il compito di promuovere l'attività e la ricerca scientifica: <https://www.nsf.gov>

³ Department of Defense Advanced Research Project Agency fu istituita nel 1958 per lo sviluppo delle tecnologie soprattutto in campo militare: <https://www.darpa.mil>

⁴ National Aeronautics and Space Administration, agenzia governativa statunitense per la ricerca aerospaziale: <https://www.nasa.gov>

⁵ La Digital Library Initiative fu una iniziativa delle tre agenzie che collaborarono dal 1994 al 1998. Nella seconda fase, dal 1998 al 2003, si unirono la National Library of Medicine, la Library of Congress e la National Endowment for the Humanities, insieme alla National Archives and Records Administration e allo Smithsonian Institution (Christine L. Borgman, 1999, p. 233).

⁶ Tra i partecipanti al gruppo di ricerca anche Marcia Bates, Efthimis Efthimiadis e Gregory Laezer. Il rapporto è disponibile a: <https://works.bepress.com/borgman/183/>

e i metadati associati; la seconda pone invece l'accento sul ruolo sociale delle biblioteche digitali, create per una comunità di utenti, nella quale i singoli interagiscono, e usano dati, informazioni e contenuti secondo i loro bisogni informativi. L'ottica, in questo caso, è l'estensione del concetto di biblioteca come istituzione e luogo fisico nel quale le risorse vengono scelte, organizzate, conservate e messe a disposizione dei lettori. La prima idea considera le biblioteche digitali come sistemi computer-based ed estensione dei sistemi di Information Retrieval; la seconda si focalizza sul fatto che esse sono una estensione dell'ambiente bibliotecario fisico e devono essere realizzate in modo che possano permettere tutte le attività che le persone fanno nel momento in cui ricercano le risorse informative.

Il workshop del 1996 ha rivestito particolare importanza perché ha messo in evidenza alcuni temi suscettibili di sviluppi nel futuro, offrendo una indicazione per le successive ricerche. Le discussioni si intrecciarono attorno a tre tematiche:

- problemi della ricerca incentrati sugli utenti, intesi sia come singoli che come comunità,
- problemi della ricerca incentrati sugli artefatti, con focus sulla creazione, organizzazione, archiviazione e ricerca dei documenti,
- problemi della ricerca incentrati sui sistemi, digital libraries come sistemi che permettono i processi di comunicazione e l'interazione con gli artefatti.

Creazione, ricerca e uso delle informazioni sono attività umane di tipo sociale, all'interno di un definito sistema sociale. La creazione dei contenuti informativi, la ricerca e la loro utilizzazione sono attività umane di tipo sociale, che comportano il continuo cambiamento e l'interazione tra i diversi attori, con reciproca influenza.

Il primo nucleo di problematiche riguarda il comportamento nella ricerca da parte delle diverse categorie di utenti – studiosi, professionisti, ragazzi – e comporta decisioni in merito alla realizzazione dei sistemi di ricerca e alla struttura delle interfacce per l'utenza nell'ambiente digitale. Già nel 1996, infatti, quel rapporto metteva in evidenza un problema fondamentale, relativo alla istituzione di una biblioteca digitale, e cioè la differenza tra una struttura progettata per sostenere la consultazione di un pubblico generale, e una struttura rivolta ad una categoria specifica, ad una comunità precisamente definita. La progettazione di una biblioteca digitale deve tenere conto delle differenze e delle variabili di tipo sociale o demografico che si possono instaurare; ma come tenere conto del fatto che gli stessi contenuti possono essere fruiti da categorie di lettori e da comunità diversi, ad esempio, contenuti digitali di tipo legale, oppure di tipo tecnico-scientifico, e come armonizzare in una biblioteca digitale le diverse necessità informative

e di studio di un'utenza eterogenea che mette in atto comportamenti di ricerca diversi?

Il secondo nucleo di problematiche offerte alla riflessione riguardava specificamente la realizzazione delle diverse tipologie di oggetti digitali, testi, immagini, suoni, dati numerici, pagine statiche o dinamiche, e dei metadati descrittivi e di tipo gestionale che li corredano. In particolare, le modalità di ricerca degli artefatti digitali possono variare a seconda delle diverse categorie di utenti, e uno dei quesiti riguardava appunto come si possano realizzare biblioteche digitali che presentino contenuti adatti a comunità diverse di utenti: come rendere fruibili i medesimi oggetti digitali specialistici, legali o medici, ad esempio, a categorie di lettori non specialisti? Un altro spunto proposto per la discussione riguardava l'indicizzazione automatica, considerata inevitabile data la massa di testi digitalizzati contenuti in una biblioteca digitale, ma al momento ritenuta insufficiente per alcune tipologie di entità.

Un terzo gruppo di riflessioni discendevano direttamente dall'insieme delle problematiche che riguardano la realizzazione degli artefatti digitali e i comportamenti di ricerca dei singoli e delle comunità, e si focalizzavano sulle tecnologie necessarie per lo sviluppo dei sistemi e sulle funzionalità offerte. L'aspetto sociale più rilevante messo in evidenza riguardava lo sviluppo delle interfacce per l'utenza e l'allestimento delle funzionalità per la ricerca adatte alle singole comunità. I nodi riguardavano la molteplicità degli utenti e le differenze di livello cognitivo e degli stadi di apprendimento raggiunti: è preferibile sviluppare interfacce diverse adatte a ciascuna categoria, oppure sviluppare un'unica interfaccia generica che si possa adattare alle diverse esigenze attraverso strumenti di navigazione appropriati? Inoltre, emergeva già nel 1996 l'esigenza di rendere interoperabili i sistemi per favorire lo scambio dei dati.

Nel complesso, il rapporto del 1996 proponeva un modello di biblioteca digitale che abbraccia due diverse visioni complementari: un modello di digital library che si propone come sviluppo dei sistemi di Information Retrieval, con attenzione agli oggetti digitali e ai loro metadati, e un modello che invece pone al centro dell'attenzione l'architettura dei sistemi e le funzionalità per la ricerca, organizzando queste ultime in modo che possano costituire il riflesso del contesto sociale entro il quale si sviluppano.

Con il National Digital Library Program (Ndlp)⁷, varato nel 1995, la Li-

⁷ <https://memory.loc.gov/ammem/dli2/html/lcndlp.html>. Questo progetto di digitalizzazione aveva l'obiettivo di offrire alla fruizione dei cittadini copie uniche di libri rari, di documenti e di fotografie rilevanti per i fini educativi, e che costituissero la memoria collettiva degli United States, ma anche di mappe e carte geografiche, di musica a stampa, di registrazioni audio e video. Si veda in proposito Caroline Arms (1996), disponibile a: <http://www.dlib.org/dlib/april96/loc/04c-arms.html>.

brary of Congress, grazie al contributo scientifico della Corporation for National Research Initiatives, aveva avviato un progetto per offrire le collezioni di documenti storici in formato digitale e renderle disponibili attraverso Internet, utilizzando anche l'esperienza fatta con il precedente American Memory Program⁸. Nel Rapporto conclusivo, pubblicato nel 1997⁹, di questo progetto coordinato da William Arms, viene presentato un prototipo dell'architettura di un sistema per la digital library e sono illustrati gli argomenti fondamentali relativi alla costruzione di una biblioteca digitale e alle sue componenti; si definisce il concetto di oggetto digitale, si forniscono informazioni sul repository e in particolare sugli identificativi persistenti degli oggetti digitali, *handles*, facendo riferimento soprattutto al fondamentale lavoro di Robert Kahn e Robert Wilensky (1995), nel quale due anni prima era stata definita l'architettura generale delle digital libraries.

Una biblioteca digitale è composta dalle risorse informative espresse in forma digitale e strutturate secondo diversi livelli di complessità, dai metadati descrittivi, strutturali, e in particolare da quelli che definiscono la proprietà intellettuale del contenuto dell'oggetto digitale. Ciascun oggetto digitale è dotato di un identificativo unico persistente, che può essere definito all'interno dell'infrastruttura e valido in quell'ambiente. Identificativi persistenti per gli oggetti digitali, per le risorse web e per i repositories sono assegnati anche nell'ambito dello *handle system*. Gli *handles* sono identificativi prodotti da strutture autorizzate a generarli, sono indipendenti dal deposito in cui gli oggetti digitali sono gestiti, sono unici, pubblici, e rimangono validi per lungo tempo¹⁰.

La biblioteca digitale si configura come una struttura complessa i cui componenti si connettono attraverso Internet: una componente basilare è il repository, il deposito in cui viene curata la gestione degli oggetti digitali e dei loro metadati e che provvede al sistema di sicurezza relativo alle operazioni fatte su ciascun oggetto. I repositories devono conservare anche i *transaction records* associati a ciascun oggetto digitale e visibili solo agli operatori autorizzati; questi registrano la data e l'ora del deposito di un og-

⁸ <http://memory.loc.gov/ammem/about/index.html>

⁹ William Y. Arms, Christophe Blanchi and Edward A. Overly (1997), <http://www.dlib.org/dlib/february97/cnri/02arms1.html>. Informazioni anche in William Y. Arms (1995), <http://www.dlib.org/dlib/July95/07arms.html> e in Varghese S. Chooralil (2010), disponibile a: <http://dspace.cusat.ac.in/jspui/bitstream/123456789/3622/1/Final.pdf>

¹⁰ La Corporation for National Research Initiatives (Cnri) gestisce il sito di *handle.net* registry (Hnr) per la distribuzione degli identificativi *handle*, e mantiene anche il Global Handle Registry (Ghr), il registro degli *handles* che sono stati attribuiti. Informazioni disponibili a: <http://www.handle.net/index.html>

getto digitale e la data e l'ora di ciascuna richiesta di recupero di ogni oggetto, insieme all'identificativo del richiedente e all'identificativo *handle*¹¹.

L'accesso al repository viene regolamentato attraverso un protocollo, Repository Access Protocol (Rap), che assicura il riconoscimento dei diritti e dei permessi d'accesso agli oggetti digitali da parte dei *client* e permette l'accesso agli oggetti digitali attraverso l'identificativo, eventualmente lo *handle*, l'accesso ai metadati, e la richiesta dell'attivazione del servizio di disseminazione delle risorse.

L'interfaccia verso l'utente si compone di due sezioni: una per gli amministratori del sistema e i bibliotecari, e una sviluppata per l'utenza esterna, la quale interagisce con il sistema tramite un browser Internet. L'interazione dell'utente con la biblioteca digitale riguarda in primo luogo i servizi sviluppati per permettere la ricerca degli oggetti digitali desiderati. I servizi del *client* provvedono a trasferire la richiesta dell'utente nei formati e con il protocollo usato dal sistema di ricerca. L'utente seleziona tra i risultati proposti dal sistema quello o quelli che ritiene rilevanti per la sua ricerca e i servizi *client*, attraverso lo *handle system* e il protocollo Rap, raggiungono il deposito, selezionano la risorsa voluta e la propongono al browser usato dall'utente.

In una biblioteca digitale gli oggetti digitali possono presentare una natura molto complessa, contenere testi strutturati in parti connesse tra loro e immagini relative ad alcune parti del testo. Le risorse informative possono presentare una gamma di varianti maggiore rispetto a quella offerta da una biblioteca cartacea: un'opera può avere diverse edizioni con modifica del testo e differenti manifestazioni a stampa, ma in ambiente digitale può presentarsi in una grande varietà di formati digitali, convertibili uno nell'altro oppure no, in una più ampia gamma di versioni, dato che l'aggiornamento dei testi è frequente e, ad esempio per le immagini, si possono avere versioni a bassa e alta risoluzione; inoltre, ciascun oggetto digitale presenta diverse restrizioni per l'uso o diverse modalità di richiesta del permesso.

L'architettura delle informazioni si fonda su di una struttura piuttosto semplice, basata sui seguenti elementi: i *data type* descrivono le proprietà tecniche dei dati, ad esempio i loro formati o i metodi usati per il loro trattamento; i metadati strutturali descrivono la tipologia degli oggetti digitali e le loro relazioni interne; i *meta-object* hanno la funzione di indici dei gruppi di oggetti digitali, dei quali in genere forniscono la lista di *handles*. Al repository nel suo insieme e ad ogni oggetto conservato sono associati i diritti e ad ogni richiesta di accesso le condizioni e i permessi per la consultazione devono essere validati.

¹¹ Per i dettagli tecnici rinvio al lavoro di Robert Kahn, Robert Wilensky (1995), disponibile a: <http://www.cnri.reston.va.us/home/cstr/arch/k-w.html>

Nel 1995 negli Stati Uniti fu istituita la Digital Library Federation (Dlf) grazie al coordinamento realizzato dalla Library of Congress e dalle biblioteche delle Università di Yale, Pennsylvania, Harvard, Stanford, Princeton, Berkeley, Michigan, Cornell, Columbia, Tennessee, cui si aggiunsero la New York Public Library e altri istituti. Oggi gli istituti coinvolti sono una quarantina, tra cui la British Library, Mit, Oxford University Library e la Bibliotheca Alexandrina. Nel 2009 la Digital Library Federation si è unita al Council on library and information resources (Clir)¹², una organizzazione indipendente che si occupa di sostenere l'accesso pubblico all'informazione e il suo mantenimento nel tempo, e alla realizzazione di servizi per la fruizione e la conservazione dei documenti, e che ha fatto propri anche i programmi che erano originariamente sostenuti dalla Dlf.

L'obiettivo che la Digital Library Federation si proponeva era curare la gestione e la conservazione di collezioni digitalizzate dei documenti cartacei fisicamente contenuti in biblioteche e archivi, e dei documenti nati digitali, che potessero documentare la cultura e la storia americane e ne offrissero ai cittadini e agli studiosi la fruizione attraverso Internet. Inoltre, si proponeva di creare una struttura per la gestione costituita da uomini e mezzi, per promuovere strategie per la creazione di altre collezioni digitali, coordinare e guidare la realizzazione di standard e buone pratiche di digitalizzazione, nonché la ricerca nell'ambito della conservazione a lungo termine degli oggetti digitali e lo sviluppo della tecnologia dedicata all'evoluzione dell'architettura dei sistemi coinvolti nell'organizzazione di biblioteche digitali.

A cura della Digital Library Federation è stato realizzato il Registry of Digital Masters, che raccoglie le informazioni relative alle monografie e alle pubblicazioni in serie nate in formato digitale o digitalizzate, in quest'ultimo caso allo scopo di evitare duplicazioni nella digitalizzazione di documenti cartacei. In Europa nel 1994 è stato costituito un consorzio tra biblioteche e reti bibliotecarie con l'obiettivo di realizzare Eromm, European Register of Microform and Digital Masters¹³, un registro pubblico che accoglie il risultato del censimento delle edizioni digitalizzate o microfilmate in modo da coordinare le informazioni relative alle digitalizzazioni per evitare inutili duplicazioni.

Nel 1998 la Digital Library Federation definiva così la biblioteca digitale:

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

¹² <http://www.clir.org/>

¹³ <http://www.eromm.org/>

digital works so that they are readily and economically available for use by a defined community or set of communities¹⁴.

In questa definizione la biblioteca digitale viene presentata come una organizzazione di risorse digitali selezionate e preservate nel tempo per la fruizione da parte di un'utenza specifica. L'accento viene posto sia sulla possibilità di far accedere gli utenti alle risorse digitali, invece che ai dati e alle informazioni, sia sull'organizzazione del servizio di conservazione nel tempo dei documenti e di utilizzazione da parte degli utenti futuri. La funzione principale di una biblioteca digitale risiede nella capacità di selezionare i documenti digitali disponibili in rete e poi catalogarli per renderli immediatamente disponibili ai lettori, oltre che nella facoltà riservata a ciascuna istituzione bibliotecaria di digitalizzare parte della propria collezione cartacea.

A seguito dell'intervento di cui si fece promotore il presidente della Repubblica francese Jacques Chirac il 28 aprile 2005, con una lettera, firmata anche dai governanti di Germania, Polonia, Italia, Spagna e Ungheria al presidente della Commissione europea José Manuel Barroso¹⁵, la Commissione europea intraprese una importante iniziativa per le biblioteche digitali: i2010 Digital libraries initiative. L'iniziativa mirava a migliorare l'usabilità delle risorse informative in rete, sia degli oggetti digitali frutto dei programmi di digitalizzazione, sia dei documenti pubblicati fin dall'origine in formato digitale. L'intervento della Commissione europea, proposto attraverso il comunicato del 30 settembre 2005 al Parlamento europeo, i2010 Digital libraries¹⁶, proponeva tre distinte linee d'azione: perfezionare le funzionalità dei sistemi e migliorare il livello di accessibilità, affinché gli utilizzatori possano trarre il massimo beneficio dalla consultazione online; incrementare i progetti di digitalizzazione di risorse cartacee, in modo da diffondere la conoscenza e favorire la società dell'informazione; rivolgere l'attenzione ai problemi della conservazione degli oggetti digitali. Oltre a concentrare la propria attenzione sul patrimonio culturale, la Commissione si occupò anche dei problemi dell'informazione scientifica e della conservazione a lungo termine delle pubblicazioni in formato digitale.

Nel dicembre 2007 l'Ifla, l'organismo internazionale che riunisce le associazioni bibliotecarie nazionali della maggior parte dei paesi del mondo,

¹⁴ www.diglib.org/about/dldefinition.htm (1998).

¹⁵ http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=5978

¹⁶ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – i2010: Digital libraries, disponibile a <https://publications.europa.eu/en/publication-detail/-/publication/e7457bac-abb7-4dd7-bc19-2c065c6b1896/language-en>

nel *Manifesto for digital libraries* ha definito la biblioteca digitale proprio come una parte fondamentale dei servizi di una biblioteca:

A digital library is an online collection of digital objects, of assured quality, [...] supported by services necessary to allow users to retrieve and exploit the resources. A digital library forms an integral part of the services of a library, applying new technology to provide access to digital collections. Within a digital library collections are created, managed and made accessible in such a way that they are readily and economically available for use by a defined community or set of communities. [...] A digital library complements initiatives to develop digital archives for the preservation of digital content. In this way, a digital library helps to improve preservation of and access to cultural and scientific heritage¹⁷.

L'Ifla ha posto l'accento sul fatto che la biblioteca digitale rende possibile l'accesso alle comunità di utenti di collezioni di documenti digitali di qualità, gestiti secondo principi condivisi a livello internazionale.

In Italia le biblioteche digitali hanno ricevuto particolare attenzione sia in ambito biblioteconomico, sia in campo archivistico. Le caratteristiche dei documenti digitali, la riproducibilità, l'autenticità e la certificazione, sono state studiate da Alberto Salarelli, *L'informazione digitale* (Salarelli, Tamaro, 2006). Questo volume è importante anche per l'inquadramento generale dell'universo delle biblioteche digitali tracciato da Anna Maria Tamaro, *La biblioteca digitale* (in particolare pp. 121-196). Le modalità della digitalizzazione delle edizioni cartacee, delle immagini e dei filmati sono state presentate da Riccardo Ridi (2004); i temi della conservazione, della certificazione e della sicurezza dei depositi, e la presentazione degli standard internazionali, sono stati materia di approfondimento in diversi contributi di studiosi del settore archivistico, e in particolare di Maria Guercio (2013). Queste tematiche, già ampiamente discusse, non sono presenti in questo volume. Questo lavoro intende infatti presentare le linee di sviluppo che le biblioteche digitali mostrano attualmente, e si concentra su alcuni aspetti, tra i quali l'applicazione delle tecnologie semantiche, come le ontologie, e sulla realizzazione dei portali per il patrimonio culturale.

¹⁷ *IFLA Manifesto for Digital Libraries*, disponibile a: www.ifla.org/publications/iflaunesco-manifesto-for-digital-libraries

Parte prima
Tipologie, modelli e funzionalità per la ricerca