

Proposta metodologica e architettuale per la gestione distribuita e condivisa di collezioni di documenti digitali

The authors address the foundations of Digital Libraries in order to envisage an innovative system architecture able to manage physically distributed archives in a conceptually integrated environment. In order to propose such a conceptual system architecture, the role of metadata for sharing information in a distributed environment is faced together with an analysis of Digital Libraries, Digital Library Systems, and Digital Library Management Systems for the management of integrated collections of digital documents that are originated in libraries, archives, and other types of production institutions. Different aspects related to the design of an archival description system in a distributed environment are considered. A matter of peculiar interest is that of metadata, so different types of metadata, their uses and sharing are presented together with the analysis of the "Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)", which is a relevant international initiative that gives directions for the management and sharing of different types of metadata in a distributed environment.

1. Introduzione

Gli archivi sono complessi di documenti, prodotti da soggetti giuridici pubblici e privati durante lo svolgimento della loro attività e rappresentano «strumento e residuo dell'attività istituzionale di un soggetto giuridico». Tale nozione di archivio non fa solo scaturire la necessità, nell'ambito archivistico, di tenere conto del contesto istituzionale e giuridico che determina la forma e la struttura dei documenti contenuti nell'archivio, ma richiede anche, avvicinandosi al problema da un punto di vista informatico e tecnologico, di preservare l'autonomia che tale contesto istituzionale e giuridico comporta.

Quando ci si appresta a progettare un sistema informatico che sia in grado di mettere in comunicazione archivi diversi al fine di facilitare la condivisione e l'accesso alle informazioni, occorre delineare un sistema informatico che non sia chiuso e monolitico ma, al contrario, sia in grado di dialogare con gli eventuali sistemi informatici di cui i diversi soggetti giuridici possono essere già forniti per la gestione del proprio archivio. D'altro canto, tale sistema informatico deve mirare a favorire l'integrazione degli archivi già esistenti e fornire un accesso

coerente a tali archivi. Ci si muove quindi tra due poli contrapposti: da un lato salvaguardare l'autonomia e le peculiarità dei vari soggetti giuridici coinvolti, dall'altro garantire che i medesimi soggetti giuridici possano beneficiare dell'integrazione e della coerenza degli archivi che partecipano al sistema informatico. Da questa dualità sorge il bisogno di ideare un sistema informatico aperto che sia in grado non solo di consentire ai soggetti giuridici che già dispongono di sistemi informatici propri, eventualmente diversi tra loro, una facile integrazione e partecipazione nel sistema complessivo, mantenendo minimali i requisiti per entrare a farvi parte, ma anche di offrire ai soggetti giuridici che ancora non dispongono di un sistema informatico proprio, la possibilità entrare nel sistema complessivo in un secondo momento e in modo incrementale.

Gli obiettivi e i requisiti appena descritti, concordano con le odierne linee di sviluppo del settore delle *digital library*. In questo contesto si preferisce mantenere la dizione inglese *digital library*, perché questo termine ha assunto un'accezione più ampia del corrispondente termine italiano "biblioteca digitale"; infatti con il termine *digital library* si indica la gestione di materiale tradizionale e digitale custodito nelle biblioteche, ma anche quello gestito negli archivi e altre istituzioni di conservazione e fruizione per il pubblico, come ad esempio i musei. Infatti le *digital library* mirano ad essere sia i depositi di varie forme di conoscenza sia i mezzi attraverso cui i cittadini possono accedere, discutere, valutare ed arricchire diversi tipi di contenuti informativi, anche con riferimento alla conservazione, alla valorizzazione e alla diffusione del patrimonio culturale¹.

In questo contesto, il ruolo dei metadati nella gestione delle collezioni di documenti è centrale e di notevole interesse, in particolare se la collezione di documenti è distribuita su più archivi. La gestione documentale influenza e contribuisce notevolmente l'accesso alle collezioni di documenti e i metadati considerati in genere utili in questo campo si rivelano allo stesso tempo di considerevole importanza anche per permettere un accesso ai documenti in base al contenuto.

¹ Y. Ioannidis – D. Maier – S. Abiteboul – P. Buneman – S. Davidson – E. A. Fox – A. Halevy – C. Knoblock – F. Rabitti – H.-J. Schek – G. Weikum, *Digital library information-technology infrastructures*, «International Journal on Digital Libraries», V/4 (2005), p. 266-274.

Il contributo di questo articolo consiste nell'analizzare e discutere la problematica della ideazione e realizzazione di un sistema informatico in grado di integrare archivi distribuiti sul territorio grazie sia ai metodi propri del settore delle digital library sia ad un opportuno scambio e condivisione di metadati tra di essi. Nell'articolo vengono discusse le principali tipologie di metadati che è necessario gestire al fine di garantire l'efficacia di un tale sistema informatico e le metodologie per la loro gestione alla luce degli standard internazionali comunemente adottati in questo contesto. Infine, viene introdotta una prima soluzione architetturale a livello concettuale per illustrare ed esemplificare il funzionamento del sistema proposto.

2. Le digital library e i sistemi di descrizione archivistica

Come discusso in precedenza, un sistema di descrizione archivistica si trova a dover soddisfare diversi obiettivi. Da un lato salvaguardare l'autonomia e le peculiarità degli archivi coinvolti e dall'altro permettere a questi archivi di beneficiare dell'integrazione e della coerenza delle informazioni condivise. Gli obiettivi di un sistema informativo di descrizione archivistica, appena considerati, concordano con le odierne linee di sviluppo del settore delle digital library.

Dal punto di vista architetturale, i sistemi più innovativi di gestione delle digital library adottano un approccio basato su servizi, cioè componenti o anche veri e propri sotto-sistemi progettati per integrarsi e cooperare insieme al fine di fornire alla digital library le funzionalità desiderate. I servizi possono interagire tra loro ed operare in modo coordinato grazie alla definizione di *interfacce*, che descrivono con precisione le funzionalità offerte da ciascun componente, e di *protocolli*, che regolano le modalità di comunicazione tra i diversi servizi e determinano il modo in cui le varie funzionalità possono essere invocate.

Dal punto di vista funzionale, i sistemi di gestione delle digital library mirano a garantire maggiore interoperabilità e integrazione tra sistemi diversi, sia a livello di metadati, che a livello di sicurezza e autenticazione, garantendo diversi gradi di visibilità e accesso alle risorse. Inoltre, tali sistemi cooperano anche nel fornire funzionalità di ricerca e di reperimento dell'informazione in risposta alle richieste di un utente, inoltrando le richieste a quei sistemi che con maggiore

probabilità sono in grado di rispondere in modo pertinente all'interrogazione.

Inoltre, le digital library cercano di offrire servizi avanzati ai propri utenti, quali annotazione dei contenuti, personalizzazione e creazione di profili di utenti, notifica e avviso quando contenuti rilevanti per l'utente vengono aggiunti o modificati, raccomandazione di ulteriori contenuti probabilmente rilevanti per l'utente, visualizzazione ed esplorazione avanzata dei contenuti o dei risultati di una ricerca, e così via.

L'elevato livello di interoperabilità è ottenuto anche grazie all'uso di contenuti informativi semi-strutturati espressi con linguaggi di *markup* estensibili, come ad esempio *eXtensible Markup Language*² (XML) definito e standardizzato dal *World Wide Web Consortium* (W3C)³, che consentono di etichettare i contenuti in modo semantico al fine di fornire sia una migliore interoperabilità sia funzionalità avanzate di ricerca dei contenuti informativi. Appare chiaro come un sistema di descrizione archivistica abbia nei suoi obiettivi funzionali aspetti comuni con le caratteristiche dei sistemi di gestione delle digital library appena descritti, soprattutto in termini di integrazione e interoperabilità, condivisione dei metadati, possibilmente espressi tramite XML, e ricerca dell'informazione di interesse. Quindi le soluzioni e le scelte proposte dalla comunità di ricerca sulle digital library possono essere prese in considerazione nel contesto della progettazione di un sistema informativo di descrizione archivistica come linee guida di riferimento.

3. Tendenze evolutive per le digital library e loro applicabilità ai sistemi di descrizione archivistica

Una digital library è un'organizzazione che ha trovato la sua forma a seguito di diversi passi di sviluppo successivi. Mentre nella metà degli anni novanta dello scorso secolo non veniva fatta distinzione fra la digital library e il sistema software incaricato della sua gestione e quindi con digital library si indicava spesso solo il sistema monolitico

² <http://www.w3.org/XML/>

³ T. BRAY – J. PAOLI – C. M. SPERBERG-MCQUEEN – E. MALER – F. YERGEAU, *Extensible Markup Language (XML) 1.0 (Fourth Edition) – W3C Recommendation 16 August 2006, edited in place 29 September 2006*, disponibile in linea all'indirizzo <http://www.w3.org/TR/REC-xml> [ultima consultazione 29 ottobre 2007].

costruito attorno alle collezioni che si dovevano gestire e alle funzionalità che si volevano offrire su tali collezioni, la tendenza odierna è invece quella di adottare un paradigma architetturale che distingue i diversi livelli di astrazione ai quali si deve operare. Inoltre, quando ci si concentra sul sistema di gestione, si preferisce considerare di dover avere a disposizione un sistema modulare e flessibile, i cui componenti possano essere riutilizzati per realizzare digital library diverse e le cui funzionalità possano essere accresciute e variate nel tempo grazie all'aggiunta o modifica dei singoli componenti che costituiscono tale sistema.

Per questa ragione, grande attenzione è posta alla realizzazione di modelli e sistemi prototipali che consentano di realizzare il paradigma descritto sopra.

In particolare, la rete di eccellenza sulle digital library DELOS⁴, finanziata nel sesto programma quadro della Comunità Europea, sta promuovendo sia la definizione di un modello di riferimento per le digital library⁵, sia lo sviluppo di un prototipo di sistema di gestione di una digital library di prossima generazione⁶, progettato in accordo con il modello di riferimento proposto.

In Figura 1 sono rappresentati graficamente i tre diversi livelli di astrazione ai quali occorre fare riferimento per abbracciare il complesso concetto di digital library; ad ogni livello corrisponde una delle

⁴ <http://www.delos.info/>

⁵ M. Agosti – L. Candela – D. Castelli – N. Ferro – Y. Ioannidis – G. Koutrika – C. Meghini – P. Pagano – S. Ross – H.-J. Schek – H. Schuldt, *A Reference Model for DLMSs Interim Report*, a cura di L. Candela e D. Castelli, *Deliverable D1.4.2 - Reference Model for Digital Library Management Systems [Draft 1]*, DELOS, *A Network of Excellence on Digital Libraries - IST-2002-2.3.1.12, Technology-enhanced Learning and Access to Cultural Heritage* (2006) – disponibile in linea all'indirizzo: http://146.48.87.122:8003/OLP/Repository/1.0/Disseminate/delos/2006_WP1_D142/content/pdf?version=1 [ultima consultazione 29 ottobre 2007].

⁶ M. Agosti – S. Berretti – G. Brettlecker – A. del Bimbo – N. Ferro – N. Fuhr – D. Keim – C.-P. Klas – T. Lidy – M. Norrie – P. Ranaldi – A. Rauber – H.-J. Schek – T. Schreck – H. Schuldt – B. Signer – M. Springmann, *DelosDLMS - the Integrated DELOS Digital Library Management System*, a cura di C. Thanos e F. Borri, *DELOS Conference 2007 Working Notes* (2007), p 71-80, ISTI-CNR, Gruppo ALI, Pisa, Italy.

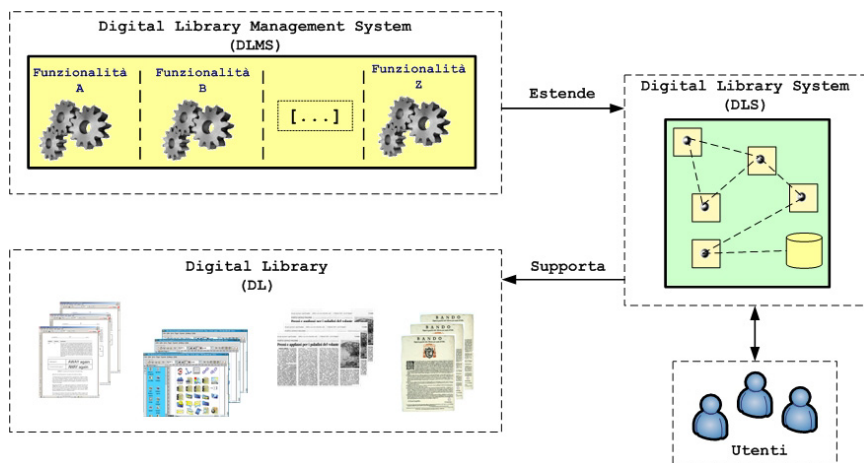


Figura 1: I tre componenti che costituiscono una “digital library”: DL, DLS e DLMS.

parti che costituisce una digital library effettiva in accordo con il modello di riferimento al quale si fa qui riferimento⁷:

- *Digital Library (DL)*: è un’organizzazione (potenzialmente virtuale) che colleziona, gestisce e conserva a lungo termine contenuti digitali e offre alla sua comunità di utenti delle funzioni specializzate sui contenuti, in misura quantificabile e in accordo con delle politiche definite.
- *Digital Library System (DLS)*: è un sistema software basato su un’architettura (potenzialmente distribuita) che provvede a tutte le funzionalità richieste da una digital library. Gli utenti interagiscono con la digital library attraverso un corrispondente digital library system.
- *Digital Library Management System (DLMS)*: è un sistema software generico che provvede ad un’appropriata architettura software che produce e amministra un digital library system. IL DLMS incorpora tutte le funzionalità che sono considerate fondamentali

⁷ L. CANDELA – D. CASTELLI – Y. IOANNIDIS – G. KOUTRIKA – P. PAGANO – S. ROSS – H.-J. SCHEK – H. SHULDY, *The Digital Library Manifesto*, in DELOS, A Network of Excellence on Digital Libraries - IST-2002-2.3.1.12, Technology-enhanced Learning and Access to Cultural Heritage (2006) – disponibile in linea all’indirizzo: http://146.48.87.122:8003/OLP/Repository/1.0/Disseminate/delos/2006_other_0081/content/pdf?version=1 [ultima consultazione 29 ottobre 2007].

per una digital library e integra anche software addizionale offrendo funzionalità più rifinite, specializzate o avanzate.

In Figura 1 possiamo vedere come questi tre componenti interagiscano tra loro. Il ruolo del DLMS risulta di evidente importanza, configurandosi come il motore di una digital library. Esso fornisce i componenti per gestire ed interagire con la digital library e può essere orientato ai servizi, risultando flessibile e quindi adattabile a diverse realtà.

I sistemi che derivano da questa organizzazione possono essere centralizzati oppure modulari. Infatti per una digital library deve esserci un solo DLS (rappresentato anche da molti piccoli DLS interconnessi tra loro) mentre tutti i DLS possono essere costituiti in molti DLMS. I servizi forniti da una digital library possono quindi essere estesi e adattati in corso d'opera e secondo le esigenze del momento. Nulla vieta di avere un sistema centralizzato e quindi in qualche modo monolitico. Un sistema di questo tipo porterà le tre componenti di una digital library a convergere in un unico componente, ottenendo una minore complessità realizzativa a prezzo di una minore adattabilità e flessibilità. D'altra parte, un sistema modulare che separa i tre componenti e vede i vari servizi offerti come dei moduli che si aggiungono al DLMS, è in genere più consono alla natura dinamica di una digital library. Un sistema di questo tipo si adatta ancora meglio e in modo naturale a una digital library distribuita. Quando si parla di digital library distribuita, non si parla solamente della distribuzione geografica delle risorse, ma anche della possibile distribuzione nell'erogazione dei servizi. Se consideriamo più sistemi partecipanti alla stessa digital library, intuiamo i vantaggi portati da un'archi-tettura modulare; ogni sistema partecipante potrebbe fornire i propri servizi in un contesto di ampia integrazione e partecipazione. Le peculiarità di ogni sistema partecipante in questo modo non sono solamente salvaguardate ma all'occorrenza possono essere sfruttate anche da altri sistemi e quindi in definitiva da una maggiore comunità di utenti.

La natura di una digital library è quindi variegata e composta da diversi componenti e attori indipendenti e interagenti tra loro. Proprio dalle modalità con cui è organizzata una digital library si intuisce come sia una realtà applicabile a diversi contesti e come si possa adattare alle necessità e particolarità di ognuno di questi. Proprio per questo motivo si ritiene che le digital library siano applicabili anche al ca-

so dei sistemi di descrizione archivistica, portando un contributo in termini di valore aggiunto per flessibilità, modularità ed estensibilità del sistema risultante.

4. Problematiche di *data curation* nei sistemi di descrizione archivistica

Un ulteriore aspetto dei moderni sistemi di gestione delle digital library, particolarmente rilevante in questo contesto, è la sempre maggiore attenzione verso alcune problematiche specifiche⁸:

- *data curation*, intesa come «l'attività di gestione e promozione dell'utilizzo dei dati, a partire dalla loro prima creazione, con la finalità di assicurare che essi si mantengano adatti agli scopi di utilizzo presente e siano disponibili per la ricerca e il riutilizzo»;
- *data archiving*, cioè «quella particolare attività della *data curation* mirata ad assicurare che i dati siano selezionati e memorizzati in maniera appropriata, possano essere accessibili, e che la loro integrità, sia logica che fisica, venga mantenuta col passare del tempo, compresa la sicurezza e l'autenticità»;
- *data preservation*, definita come «quella particolare attività del *data archiving* in cui specifici dati vengono mantenuti nel tempo in modo che siano possibili l'accesso e la comprensione anche a seguito di cambiamenti nelle tecnologie».

Parte dell'attività di data curation riguarda dunque la descrizione delle origini dei dati e il tenere traccia del processo secondo cui tali dati sono giunti in un archivio e delle modifiche che hanno subito nel corso del tempo. Tale attività risulta essere fondamentale per poter affrontare in modo scientifico i contenuti informativi gestiti da un archivio, in quanto offre la possibilità di ricostruire il percorso di un elemento d'informazione per poter essere in grado di interpretarlo correttamente alla luce del contesto in cui è stato creato e della sua evoluzione⁹.

⁸ P. LORD – A. MACDONALD, e-Science Curation Report. Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision, The JISC Committee for the Support of Research, 2003, disponibile in linea all'indirizzo: http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal. [ultima consultazione, 29 ottobre 2007].

⁹ S. Abiteboul – R. Agrawal – P. Bernstein – M. Carey – S. Ceri – B. Croft – D. De Witt – M. Franklin – H. Garcia-Molina – D. Gawlick – J. Gray – L. Haas – A.

Risulta quindi chiaro, alla luce della discussione già condotta, come le tematiche della data curation siano di fondamentale importanza nell'ambito archivistico, sia al fine di preservare e tenere traccia del contesto istituzionale e giuridico, che caratterizza i documenti di un archivio, sia al fine di evitare la progettazione e lo sviluppo di un sistema informatico che appiattisca e che impedisca di evidenziare come la descrizione archivistica abbia connotati storiografici e amministrativi connessi al periodo in cui si è sviluppata.

D'altro canto, il conseguimento di livelli di data curation elevati e automatizzati all'interno di un sistema di gestione di digital library viene usualmente indicato come un obiettivo a medio-lungo termine. Tale requisito rimane comunque fondamentale per un sistema informatico che abbia gli obiettivi generali descritti in precedenza.

Occorre dunque tenere conto di questo requisito sin dalle prime fasi di analisi e progettazione di un sistema, in modo da garantire che l'architettura di tale sistema sia sufficientemente aperta ed estensibile da poter supportare l'introduzione di funzionalità avanzate di data curation in un secondo momento, delineando un percorso che consenta lo sviluppo progressivo e incrementale del sistema stesso. Inoltre un'architettura modulare ed aperta consente la definizione di diversi livelli di integrazione di sistemi di vari soggetti giuridici. È possibile, ad esempio, fissare un *livello di integrazione di base*, che consiste nella possibilità di condividere metadati e descrizioni archivistiche, mentre un *livello di integrazione avanzato* può essere caratterizzato da interoperabilità nelle funzionalità di data curation. In questo modo è possibile adottare un approccio progressivo e differenziato nella partecipazione di diversi archivi; infatti alcuni soggetti giuridici potrebbero partecipare solo a livello di integrazione di base, mentre altri soggetti potrebbero adottare il livello di integrazione avanzato.

Il beneficio di questo approccio è far sì che lo sforzo necessario perché un soggetto giuridico possa prendere parte al sistema informatico, sia in termini di risorse umane e informatiche che in termini

Halevy – J. Hellerstein – Y. Ioannidis – M. Kersten – M. Pazzani – M. Lesk – D. Maier – J. Naughton – H.-J. Schek – T. Sellis – A. Silberschatz – M. Stonebraker – R. Snodgrass – J. D. Ullman – G. Weikum – J. Widom – S. Zdonik, *The Lowell Database Research Self-Assessment*, «Communications of the ACM (CACM)», XLVIII /5 (2005), p. 111–118.

di risorse economiche, sia modulabile sulle esigenze e sulle caratteristiche di tale soggetto giuridico.

5. Ruolo e tipologie di metadati

In molti sistemi informativi è ormai consueto avere a che fare con collezioni di documenti distribuite in più depositi – normalmente denominati con il termine inglese *repository* – vista la grande mole di informazioni che devono essere gestite e recuperate, considerato in questo senso anche l’apporto delle collezioni di documenti accessibili grazie al Web.

In genere la tendenza è quella di distribuire i documenti su più repository per facilitarne la gestione e la conservazione. Invero grandi repository monolitici permettono di avere tutti i documenti nel medesimo spazio fisico, ma richiedono un gran numero di risorse per la conservazione e forniscono dei servizi sui documenti la cui efficienza tende a diminuire al crescere delle dimensioni dei repository stessi.

Una collezione di documenti può essere rappresentata da un unico insieme logico pur non essendo contenuta in un unico repository. Si può quindi separare la gestione logica dei documenti da quella fisica, in modo da poter sfruttare tutti i vantaggi dell’aver più repository distribuiti. La distribuzione dei documenti può avvenire su più repository distanti tra loro geograficamente e che per condividere i propri contenuti informativi devono servirsi della rete Internet. In questo contesto si devono poter gestire e scambiare i documenti in modo definito da procedure condivise e standardizzate.

Una situazione di questo tipo è rappresentata da un sistema di descrizione archivistica che, come scopo, si pone quello di permettere la condivisione dell’informazione presente in documenti archivistici conservati in vari archivi distanti geograficamente tra loro. In questo contesto è necessario gestire e rendere accessibili le descrizioni archivistiche mantenendo l’indipendenza di ogni singolo archivio. Si deve prevedere di gestire le descrizioni archivistiche a livello logico in modo indipendente dalla gestione dei documenti archivistici, per la cui descrizione si utilizzano dei metadati. I *metadati* nel mondo delle digital library sono definiti come “dati sui dati”, perché sono dei dati che permettono di descrivere altri dati e che vengono utilizzati come

descrizioni strutturate e standardizzate delle risorse presenti in una digital library.

In un sistema di descrizione archivistica distribuito, si dovranno valutare principalmente due tipi di metadati.

Il primo tipo di metadati è quello che permette di descrivere i documenti detenuti da ciascun archivio. Questi metadati, che possono essere chiamati *metadati descrittivi*, permettono di ottenere informazioni sui documenti ai quali fanno riferimento e vengono gestiti in maniera indipendente dai documenti stessi. I metadati descrittivi permettono di descrivere i documenti d'archivio senza avere né la funzione né la ragione di essere anche solo un "compendio" di questi. I metadati non portano un'informazione surrogata o sintetizzata di quella che si può ricavare dai documenti, ma una vera e propria "altra informazione" legata all'identificazione, localizzazione e reperimento dei documenti. I metadati descrittivi, una volta condivisi, potranno dare una visione complessiva dei documenti di ogni archivio locale, mentre la gestione dei documenti veri e propri verrà lasciata agli archivi che li detengono. Un sistema di descrizione archivistica fornirà agli utenti dei servizi che si baseranno fundamentalmente sui metadati descrittivi; ad esempio la ricerca di un particolare documento verrà eseguita sui metadati e non direttamente sui documenti detenuti dall'archivio di competenza.

Il secondo tipo di metadati che si deve considerare per un sistema di descrizione archivistica è collegato alla formazione e gestione di un *archivio* o *lista di autorità* – in inglese *authority file*. L'archivio o lista d'autorità è un archivio di intestazioni che costituiscono un punto d'accesso effettivo o potenziale ai contenuti di un archivio; l'archivio d'autorità è un dizionario dei valori di quel tipo d'accesso. L'archivio d'autorità può essere costituito a priori e senza collegamenti ad oggetti fisici, quindi è un archivio che può essere costruito preventivamente utilizzando la forma normalizzata di tutti gli elementi che costituiscono un accesso potenziale d'interesse. La funzione di normalizzazione può essere esaltata inserendo nel dizionario anche tutte le varianti utili delle intestazioni, accompagnate dal rinvio esplicito alla rispettiva forma ufficiale.

L'ideazione e la realizzazione di un archivio d'autorità richiede che venga esercitato un controllo sulle intestazioni che possono esse-

re utilizzate, per questo si parla della necessità di un *controllo d'autorità* – in inglese *authority control* – quando si costruisce un archivio, che possa essere effettivamente considerato di “autorità” per quello specifico tipo di accesso ai contenuti dell’archivio. In italiano il termine inglese *authority* viene normalmente reso con il termine “autorità”, ma è necessario tenere presente che, quando si progetta e si concretizza un archivio o lista d’autorità, si fa un lavoro che richiede l’esercizio della “autorevolezza” e non solo quello dell’autorità. Infatti, il processo di costruzione di un archivio d’autorità richiede lo studio delle consuetudini dei professionisti e degli utenti finali che operano nello specifico contesto di riferimento, quindi, la corrispondente scelta delle intestazioni deve essere fatta in modo ponderato e autorevole. Qui si adottano i termini lista e archivio d’autorità, perché sono quelli in uso più comune, anche se è sembrato utile richiamare la sfumatura d’uso del termine per contestualizzarne meglio la finalità.

L’importanza di tenere presente, nel progetto e nella realizzazione di in un sistema di gestione automatica delle informazioni bibliografiche, la gestione automatica delle liste o archivi d’autorità, per ogni tipo di accesso considerato utile e necessario per l’utenza, è stata messa in evidenza da tempo¹⁰; in anni recenti l’importanza degli archivi d’autorità è apparsa di rilievo per molti anche nella progettazione di nuovi sistemi di digital library.

Ogni archivio o lista d’autorità permette di definire univocamente i valori di un tipo d’accesso, quindi i valori che vengono associati ad un’entità, sia essa una persona, un’organizzazione o un’istituzione d’interesse. Il mantenimento, la gestione e l’aggiornamento di ogni lista è fondamentale, soprattutto in un sistema distribuito, per mantenere i dati aggiornati in base all’uso effettivo che ne viene fatto nel settore di riferimento.

Se si considera un sistema a “mondo chiuso” la definizione delle liste d’autorità, e quindi il problema dell’*authority control*, è di minor complessità, in quanto è richiesto che queste liste siano consistenti all’interno di un sistema in genere controllabile e di medie/piccole dimensioni. Se consideriamo invece un insieme di archivi distribuito, tenendo conto anche della sua possibile apertura nel Web, la dimen-

¹⁰ M.B. BALDACCI – R. SPRUGNOLI, *Informatica e biblioteche*, Roma, La Nuova Italia Scientifica, 1983.

sione del problema aumenta in modo considerevole superando non solo i confini definiti dall'archivio di un ente ma anche quelli di una singola nazione, rendendo estremamente difficile il controllo delle risorse presenti e quindi anche la consistenza delle liste d'autorità. Il problema può essere affrontato definendo le liste d'autorità come un particolare tipo di metadato.

Ci troviamo quindi di fronte ad una realtà che deve gestire i metadati descrittivi formati sui documenti d'archivio e i *metadati di autorità* che costituiscono le liste d'autorità. Ogni archivio deve quindi considerare due tipi diversi di metadati e dalla descrizione appena fatta si evince come la loro gestione richieda uno sforzo diversificato. Infatti, i metadati descrittivi sono creati dall'archivio stesso sui documenti da esso detenuti, mentre i metadati d'autorità potranno essere formati da un'organizzazione centrale e autorevole e poi condivisi da tutti gli archivi partecipanti al servizio di descrizione.

Il vantaggio costituito dall'uso dei metadati deriva dalla loro natura di oggetti strutturati e allo stesso tempo flessibili. Il formato e il tipo di informazioni che un metadato deve contenere sono specificati da regole ben definite che permettono di stabilirne con accuratezza la composizione. Ogni tipo di metadato è quindi formato secondo regole precise e formalizzate in un archivio digitale – in inglese *file* – identificato come *schema del metadato*. Conoscendo lo schema di un metadato sapremo a priori il formato e il tipo di informazioni portate da quel particolare tipo di metadato. Ne consegue che i metadati garantiscono un'informazione strutturata su campi ben definiti e allo stesso tempo la definizione di diversi schemi di metadato permette di differenziare il tipo di informazioni che di volta in volta si intendono fornire in relazione alle risorse descritte. Inoltre, i metadati sono gestiti con file di dimensioni ridotte e quindi un sistema basato su di essi può risultare flessibile e scalabile.

Queste caratteristiche ben si adattano al sistema di descrizione archivistica che si è finora ipotizzato. Si può prevedere che un sistema come questo operi effettuando uno scambio di metadati descrittivi e di metadati d'autorità tra i vari archivi partecipanti. Ogni archivio può condividere i propri metadati descrittivi con tutti gli altri archivi e ognuno di questi può recuperare i metadati d'autorità da un repository centrale. La definizione di un sistema di descrizione archivistica

basato sull'uso dei metadati permette di definirne il funzionamento a livello concettuale senza dover scegliere una particolare piattaforma tecnologica. Infatti uno schema di metadati è indipendente da specifiche implementazioni e favorisce l'interoperabilità tra sistemi diversi.

La tendenza degli odierni sistemi informativi e dei protocolli di scambio di metadati è quella di garantire un funzionamento il più indipendente possibile dalla scelta di un particolare schema di metadati. Questo significa che i metadati descrittivi e i metadati d'autorità, pur essendo tra loro differenti, possono essere gestiti dai medesimi sistemi informativi e possono essere trasportati mediante gli stessi protocolli. Sviluppando un sistema di descrizione archivistica che utilizza i metadati come mezzo di scambio di informazioni si può avere un insieme di archivi che sono totalmente indipendenti l'uno dall'altro, ma in comunicazione tra loro perché possono condividere i propri contenuti informativi attraverso i metadati.

6. Il protocollo OAI-PMH per lo scambio dei metadati

Nell'ambito della condivisione e del trasporto dei metadati si introduce ora un protocollo che permette di affrontare e risolvere il problema di rappresentazione e gestione dei metadati e che si dimostra particolarmente adattabile alle digital library. Il protocollo di scambio metadati in questione è l'*Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH)*¹¹, sviluppato nell'ambito dell'iniziativa internazionale denominata *Open Archives Initiative (OAI)*¹². La OAI si propone di sviluppare e promuovere degli standard per favorire l'interoperabilità tra diversi sistemi informatici in modo da permettere la condivisione di informazioni, principalmente mediante lo scambio di metadati fra i sistemi stessi.

L'uso della parola *archives*, nel contesto della OAI, non deve trarre in inganno, in quanto non deve essere considerata in senso stretto, ma nella più ampia accezione del termine. Con *archives* si vuole indica-

¹¹ H. VAN DE SOMPEL – C. LAGOZE – M. L. NELSON – S. WARNER, *The Open Archives Initiative Protocol for Metadata Harvesting (2nd ed.)*, Open Archives Initiative, p. 24, 2003, disponibile in linea all'indirizzo: <http://www.openarchives.org/OAI/openarchi-vesprotocol.html> [ultima consultazione 29 ottobre 2007].

¹² <http://www.openarchives.org/>

re un qualsiasi repository documentale senza che questo debba essere per forza un archivio nel senso tecnico del termine.

OAI-PMH stabilisce una serie di regole per la comunicazione tra sistemi, formalizzando un meccanismo per la raccolta dei metadati da vari repository. Il punto nodale di questo protocollo, e ciò che principalmente lo differenzia da altre soluzioni, è da ricercare nel significato della parola “*harvesting*” (raccogliere, mietere). Infatti OAI-PMH non è stato progettato per la realizzazione di sistemi che effettuano una ricerca distribuita su repository, ma ipotizza la raccolta di metadati in uno o più repository centralizzati. Dall’esperienza delle digital library si è verificato che la ricerca distribuita si dimostra lenta e sempre meno affidabile al crescere del numero di repository considerati.

La pratica promossa da OAI-PMH, tesa alla centralizzazione dei servizi, resa possibile dalle caratteristiche intrinseche dei metadati, invece si dimostra scalabile e particolarmente efficiente anche all’aumentare del numero dei repository considerati¹³. Si devono infatti tener presente le semplificazioni introdotte da OAI-PMH. Se pensiamo ad un sistema che realizza una ricerca su più repository distribuiti in rete, quando aumenta il numero dei repository aumenta il numero dei sistemi presenti in nodi della rete sui quali realizzare la ricerca, aggiungendo un crescente grado di complessità al sistema che va a peggiorare l’efficienza e l’efficacia del meccanismo di ricerca. Fornendo, invece, dei servizi su un insieme di metadati centralizzato, il numero dei nodi sui quali realizzarli viene ridotto a uno mantenendo i servizi indipendenti dalle numero di repository e di documenti considerati.

È sempre da tenere in considerazione che OAI-PMH è un protocollo di scambio e quindi permette di raccogliere i metadati ma non realizza alcun servizio su questi. Per fornire servizi sui metadati, il protocollo deve essere abbinato ad altre soluzioni.

Il funzionamento del protocollo si fonda su due componenti principali chiamate *data provider* e *service provider*, rispettivamente fornitore di risorse (metadati) e di servizi. Il componente data provider mantiene le risorse ed espone i propri metadati per la raccolta. In un sistema distribuito, il concetto di fornitore può essere esteso e si può

¹³ S. L. SHREEVES – T. G. HABING – K. HAGEDORN – L. A. YOUNG, *Current Developments and Future Trends for the OAI Protocol for Metadata Harvesting*, «Library Trends» 53, 4 (2005), p. 576-589.

pensare di abbinare ad ogni repository un data provider che fornisce metadati nell'ambito di OAI-PMH. I data provider sono quindi i soggetti produttori di metadati e i conservatori delle risorse. I service provider, invece, raccolgono i metadati dai vari data provider e forniscono su questi dei servizi, come ad esempio delle funzionalità di ricerca.

I data provider devono dotarsi di un'interfaccia per rispondere alle richieste dei service provider, ogni interfaccia può essere utilizzata mediante un software appositamente sviluppato chiamato *harvester*. Il software harvester viene utilizzato per interrogare, secondo delle specifiche richieste definite dal protocollo, i vari data provider ottenendo in risposta i metadati richiesti. Un sistema basato su OAI-PMH è composto almeno da un service provider e da due data provider ai quali il fornitore di servizi chiede i metadati, ma nulla vieta che ci sia la presenza di più service provider.

La Figura 2 permette di avere una visione complessiva dell'architettura concettuale di un possibile sistema distribuito basato sul protocollo OAI-PMH. Ogni repository si interfaccia con un data provider che fornisce i metadati, costruiti sulle risorse dei repository, ai service provider che li richiedono. I service provider raccolgono i metadati mediante il software harvester e ne costituiscono un insieme centralizzato, sul quale vengono forniti dei servizi utilizzabili dagli utenti finali. Come si può vedere in Figura 2, più service provider possono interrogare lo stesso data provider.

OAI-PMH, essendo un protocollo teso a favorire l'interoperabilità tra diversi sistemi, permette di utilizzare molteplici formati di metadati definiti secondo diversi schemi. In questo modo OAI-PMH può essere utilizzato per lo scambio di metadati di tipo diverso. Considerando un sistema di descrizione archivistica, si può prevedere di utilizzare il protocollo per scambiare in modo omogeneo sia i metadati descrittivi che i metadati di autorità. Questa caratteristica permette di adattare un sistema che utilizza OAI-PMH a possibili nuovi schemi di metadati che potranno essere sviluppati nel tempo grazie alla continua innovazione tecnologica a cui stiamo assistendo.

In questo modo un sistema di questo tipo è affetto in maniera meno pesante, rispetto ad altri meno flessibili, dai problemi legati all'obsolescenza tecnologica.

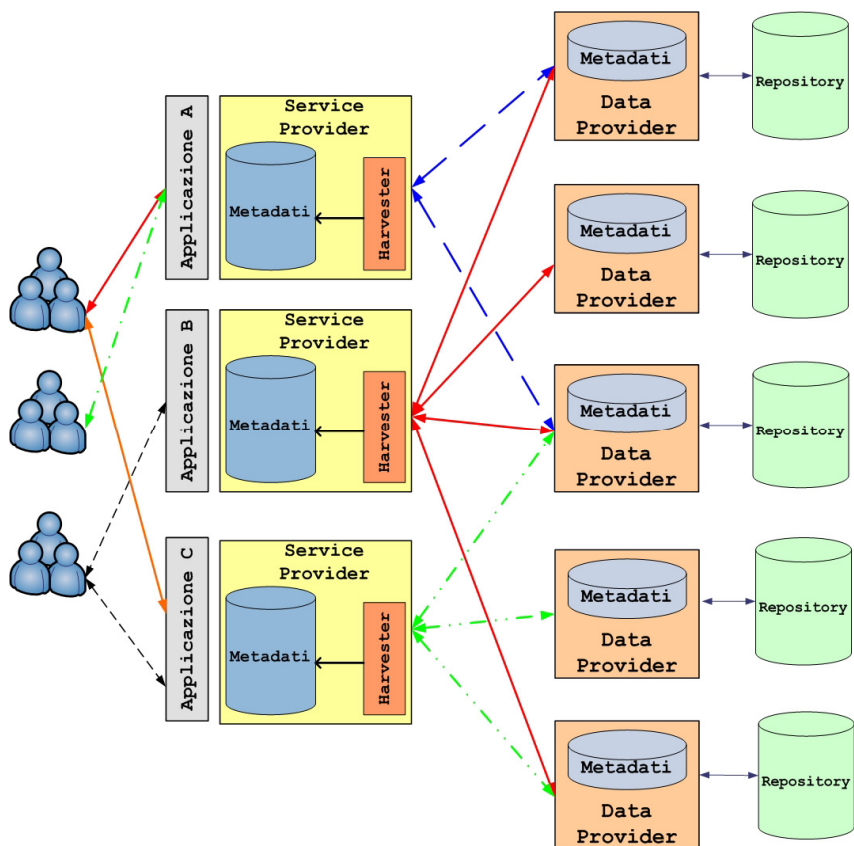


Figura 2 Architettura concettuale OAI-PMH: diversi repository scambiano i propri documenti con differenti data provider per permettere la formazione e la raccolta dei metadati da parte dei service provider.

Il protocollo OAI-PMH permette la raccolta incrementale dei metadati dai vari repository. Questo significa che ogni volta che un service provider interroga un repository non raccoglie ogni volta tutti i metadati presenti nel data provider, ma solamente quelli che non possiede già. Questo è possibile grazie all'implementazione di una funzione che permette di verificare la data in cui un metadato è stato inserito in un data provider e di raccogliarlo solo se la suddetta data è posteriore a quella dell'ultima visita del software harvester. Chiaramente il service provider dovrà visitare periodicamente tutti i data provider dai quali raccoglie metadati, stabilendo una strategia che gli

permetta di non sprecare risorse visitandone troppo spesso di non aggiornati e allo stesso tempo di non lasciar passare troppo tempo senza aggiornare la propria raccolta di metadati. Non è possibile stabilire a priori quale sia la miglior strategia di raccolta implementabile perché questa dipende fortemente dal tipo di documenti gestiti dai repository. Supponiamo infatti di dover raccogliere i metadati forniti da un archivio storico, in questo caso una frequenza di raccolta giornaliera è probabilmente eccessiva se si tiene presente la cadenza degli aggiornamenti dell'archivio. Se consideriamo, invece, un repository contenente le notizie di un quotidiano on-line, una frequenza di raccolta giornaliera sarà del tutto insufficiente. Per questo motivo i parametri che stabiliscono le politiche di aggiornamento e uso delle funzionalità del OAI-PMH dovranno essere inizializzati e stabiliti di volta in volta dopo un'attenta analisi dello specifico sistema che si sta realizzando¹⁴.

La questione relativa alla frequenza di raccolta dei metadati e quindi di aggiornamento dei repository centralizzati, deve essere valutata a maggior ragione quando si considerano i metadati d'autorità, in modo da evitare di avere liste di autorità in repository centrali inconsistenti con quelle presenti nei repository locali.

7. Una significativa esperienza d'integrazione di sistemi diversi basata sui metadati

A livello europeo nel settore delle digital library, un'interessante esperienza di integrazione, del tipo discusso in precedenza, è rappresentata da *The European Library*¹⁵, un servizio, accessibile tramite un portale Web, che ambisce a fornire un'infrastruttura per offrire un punto di accesso comune e integrato alle collezioni delle biblioteche nazionali degli stati membri della Comunità Europea. Il sistema offerto da *The European Library* è considerato uno dei componenti fon-

¹⁴ X. LIU – K. MALI – M. ZUBAIR – M. L. NELSON, *Repository synchronization in the OAI framework*, a cura di L. DELCAMBRE – G. HENRY – C. C. MARSHAL, Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries, p. 191-198, IEEE Computer Society, Washington, DC, USA, 2003.

¹⁵ <http://www.theeuropeanlibrary.org/>

damentali per la *i2010 Digital Library Initiative*¹⁶, avviata nel contesto del settimo programma quadro della Comunità Europea, e che mira a sviluppare una digital library comune a livello europeo per l'anno 2010.

Attualmente *The European Library* realizza la federazione di 32 biblioteche nazionali per un totale di 307 collezioni disponibili e accessibili per l'utente finale. Il numero di biblioteche nazionali che entrano a far parte del sistema è in costante crescita in quanto l'obiettivo finale è integrare le biblioteche nazionali di tutti gli stati membri.

Vi sono varie modalità secondo cui una biblioteca nazionale può entrare a far parte del sistema *The European Library* a seconda delle caratteristiche del sistema informatico in uso nella biblioteca. In ogni caso, l'approccio di *The European Library* è mantenere minimi i requisiti per far sì che una biblioteca nazionale entri a far parte del sistema¹⁷, di modo che ogni biblioteca nazionale possa preservare la propria autonomia, mantenere la gestione e l'accesso alle proprie collezioni e non sia costretta a modificare i propri sistemi informatici.

Una delle modalità possibili è basata sul protocollo OAI-PMH che consente di raccogliere i record catalografici presenti nelle varie biblioteche nazionali; in questa configurazione il sistema *The European Library* agisce come *service provider* e raccoglie le descrizioni bibliografiche dai sistemi informatici delle varie biblioteche nazionali che hanno il ruolo di *data provider*. Le informazioni e i metadati raccolti vengono memorizzati in un indice centralizzato, che consente poi di offrire funzionalità di ricerca avanzata sulle informazioni in esso memorizzate. Inoltre, a partire dalle informazioni contenute nell'indice centralizzato, sono in fase di sviluppo nuovi servizi, come ad esempio servizi di ricerca multilingue capaci di recuperare documenti di diverse lingue in risposta ad un'interrogazione dell'utente¹⁸.

¹⁶ EUROPEAN COMMISSION INFORMATION SOCIETY AND MEDIA, *i2010: Digital Libraries*, ottobre 2006 – disponibile in linea all'indirizzo: http://europa.eu.int/information_society/activities/digital_libraries/doc/brochures/dl_brochure_2006.pdf [ultima consultazione 29 ottobre 2007].

¹⁷ T. VAN VEEN – B. OLDROYD, Search and Retrieval in The European Library. A New Approach. «D-Lib Magazine», X/2 (2004).

¹⁸ M. AGOSTI – M. BRASCHLER – N. FERRO – C. PETERS – S. SIEBINGA, *Roadmap for MultiLingual Information Access in The European Library*, a cura di N. FUHR – L. KOVACS – C. MEGHINI, Proc. 11th European Conference on Research and Ad-

L'esempio della soluzione attuata nel sistema *The European Library* mostra che è possibile seguire la strada dell'integrazione di sistemi diversi basandosi sullo scambio di metadati tramite il protocollo OAI-PMH; infatti la soluzione proposta ha dimostrato che è possibile integrare sistemi eterogenei tra loro preservando le rispettive autonomie, in una soluzione effettivamente scalabile; la soluzione proposta, poi, può costituire il punto di partenza per sviluppare servizi avanzati. Inoltre, l'esempio dimostra una volta ancora come possa essere di beneficio prendere in considerazione le tecniche adottate nel settore delle *digital library* per sviluppare un sistema di gestione di descrizioni archivistiche.

A questo proposito è utile porre in evidenza come l'approccio adottato per *The European Library* venga utilizzato anche in contesti diversi da quello delle biblioteche nazionali. È questo il caso della rete tematica EDLnet¹⁹, avviata sempre nel contesto dell'iniziativa i2010, che mira a sviluppare soluzioni per l'interoperabilità tra i sistemi che gestiscono il patrimonio culturale europeo nei musei, negli archivi, negli archivi di materiale audio-visivo e nelle biblioteche. L'integrazione dei sistemi di queste istituzioni diverse si baserà sull'esperienza e le strategie adottate nel sistema *The European Library* discusso in precedenza.

8. Proposta di architettura concettuale

Da quanto si è analizzato finora emerge come un sistema distribuito di descrizione archivistica si possa attuare secondo le modalità e in qualche modo la filosofia ispirata al mondo delle digital library. L'utilizzo delle metodologie e tecnologie sviluppate nel settore delle digital library, permette di affrontare il problema della distribuzione delle collezioni, quindi del sistema di descrizione e anche di quello del controllo d'autorità. Come abbiamo visto le liste o archivi d'autorità possono essere gestiti come metadati, con le modalità che ad essi appartengono. Certamente gli archivi d'autorità introducono un livello di complessità maggiore, si è infatti parlato della necessità di formare

vanced Technology for Digital Libraries (ECDL 2007), p. 136-147, Lecture Notes in Computer Science (LNCS) 4675, Springer, Heidelberg, Germany, 2007.

¹⁹ <http://www.europeandigitallibrary.eu/edlnet/>

un'autorità centrale per ogni insieme di collezioni che in qualche modo regolamenti la creazione, l'aggiornamento e la diffusione di questi archivi.

L'analisi condotta ci porta a introdurre un protocollo per lo scambio di metadati che può essere applicato con consistenza alle digital library e quindi anche ad un sistema di descrizione archivistica, permettendo di affrontare anche la problematica del controllo d'autorità. In un contesto di descrizione archivistica è pensabile di avere molti *data provider* e un *service provider*. Ossia molti produttori di documenti che mettono a disposizione i propri metadati e un raccoglitore che interroga i vari produttori per ottenere i metadati descrittivi. Il *service provider*, poi, deve mettere a disposizione degli utenti degli strumenti di ricerca per accedere ai contenuti. In questo modo si può pensare una architettura realizzativa che permette di accedere centralmente, e quindi da un unico punto d'accesso, ad una moltitudine di collezioni ampliando notevolmente il patrimonio informativo, senza intaccare l'indipendenza e le politiche di gestione di ogni singolo produttore. Mediante OAI-PMH, quindi, non vengono spostate le risorse vere e proprie ma i metadati che le descrivono.

Se consideriamo gli archivi di autorità la questione potrebbe essere simmetrica, ossia avremo un unico *data provider* che fornisce gli archivi d'autorità e molti *service provider* che li raccolgono. In questo modo perseguiamo l'idea di avere una sorta di organismo centrale che funge da componente di authority control, garantendo la creazione, il mantenimento e l'aggiornamento degli archivi d'autorità. Si evita di avere così un insieme di archivi d'autorità afferenti alla stessa entità, in contrasto tra loro.

Se dovessimo pensare ad un'architettura che unisce questi due aspetti²⁰, da un lato il sistema di descrizione archivistica e dall'altro la gestione degli archivi d'autorità, vedremmo che ogni organismo avente un *data provider* nel primo caso dovrebbe dotarsi anche di un *service provider* nel secondo e viceversa.

²⁰ M. AGOSTI – N. FERRO – G. SILVELLO, *An Architecture to Share Metadata among Geographically Distributed Archives*, a cura di C. THANOS – F. BORRI, *DELOS Conference 2007 Working Notes* (2007), p. 231-239, ISTI-CNR, Gruppo ALI, Pisa, Italy.

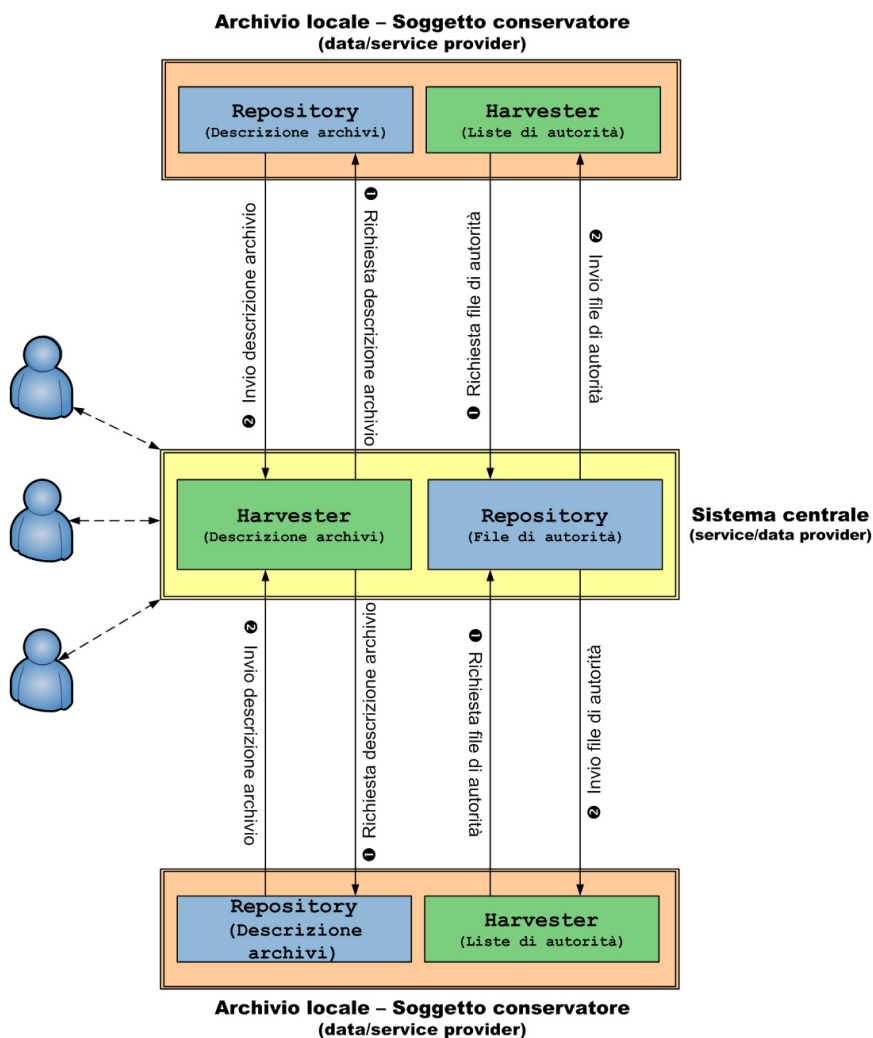


Figura 3: Architettura concettuale per un sistema di descrizione archivistica.

La Figura 3 permette di visualizzare questo concetto, presentando la possibile architettura concettuale per il sistema appena descritto; come si vede, l'architettura si basa su quanto previsto da OAI-PMH. L'architettura poi prevede che l'accesso da parte dell'utente, sia esso un utente fruitore o un utente gestore/archivista, avvenga sempre tramite gli strumenti di gestione e applicativi del Web; quindi

L'accesso è possibile tramite una qualsiasi postazione connessa alla rete Internet e dotata di un programma di accesso al Web, programma che viene normalmente denominato *browser Web*, perché permette la consultazione delle risorse informative accessibili via Web.

Come si può osservare dalla Figura 3, l'architettura proposta presenta una notevole simmetria: il sistema centrale si occupa di gestire e mantenere in modo uniforme le liste o file d'autorità, agendo come data provider verso gli archivi locali che usano tali liste d'autorità per creare le proprie descrizioni archivistiche. In maniera speculare, gli archivi locali agiscono come data provider nei confronti del sistema centrale. Il sistema centrale si occupa di fornire servizi avanzati ai propri utenti tramite le descrizioni archivistiche raccolte dai singoli archivi locali.

Nel seguito, si conclude l'illustrazione della soluzione proposta con la presentazione di due possibili scenari di utilizzo dell'architettura stessa, perché in questo modo è possibile verificare quale sia l'effettiva operatività della soluzione proposta.

Esempio di utilizzo da parte di un utente fruitore

Questo esempio fa riferimento ad uno scenario in cui un utente fruitore accede al sito Web del sistema centrale. In primo luogo, si ipotizza che l'utente consulti le liste d'autorità per individuare un elemento di suo interesse; in secondo luogo, si ipotizza che esso cerchi nell'archivio di un soggetto conservatore maggiori informazioni che riguardano l'elemento di suo interesse; infine, che esso consulti le informazioni messe a disposizione direttamente dal soggetto conservatore. In particolare:

- l'utente fruitore richiede, tramite il browser Web, la lista di autorità gestita dal sistema centrale;
- il modulo di interfaccia utente del sistema centrale inoltra la richiesta al modulo di gestione delle liste d'autorità. Il modulo di gestione delle liste d'autorità restituisce la lista d'autorità al modulo di interfaccia utente che, a sua volta, la restituisce al browser Web per la presentazione all'utente;
- l'utente consulta nel browser Web la lista d'autorità ed individua un elemento di suo interesse. A questo punto richiede la descri-

zione degli elementi dell'archivio di un soggetto conservatore che sono pertinenti all'elemento di suo interesse;

- il modulo di interfaccia utente del sistema centrale inoltra la richiesta al modulo di acquisizione delle descrizioni degli archivi. Il modulo di acquisizione delle descrizioni degli archivi restituisce le descrizioni pertinenti al modulo di interfaccia utente che, a sua volta, le restituisce al browser Web per la presentazione all'utente;
- l'utente consulta attraverso il browser Web le descrizioni ed individua un elemento di suo interesse. A questo punto richiede l'elemento di suo interesse direttamente all'archivio di un soggetto conservatore;
- il modulo di interfaccia utente del soggetto conservatore inoltra la richiesta al modulo di gestione delle descrizioni degli archivi. Il modulo di gestione delle descrizioni degli archivi restituisce le descrizioni pertinenti al modulo di interfaccia utente che, a sua volta, le restituisce al browser Web per la presentazione all'utente;
- l'utente consulta l'elemento di suo interesse e conclude questa fase di interazione con il sistema.

Si noti come parte dell'interazione avvenga tra l'utente e il sistema centrale e parte dell'interazione avvenga direttamente tra l'utente e il sistema del soggetto conservatore.

Esempio di utilizzo da parte di un utente archivista

Questo esempio fa riferimento ad uno scenario in cui un utente gestore/archivista accede al sito Web del sistema ospitato dal proprio soggetto conservatore. In primo luogo, l'utente consulta le liste d'autorità per individuare un elemento di suo interesse; in secondo luogo, recupera dal sistema centrale maggiori informazioni relative all'elemento di suo interesse; infine, utilizza tali informazioni per creare una nuova descrizione nell'archivio del proprio soggetto conservatore. In particolare:

- l'utente fruitore richiede, tramite il browser Web, la lista di autorità acquisita dal sistema del soggetto conservatore;
- il modulo di interfaccia utente del soggetto conservatore inoltra la richiesta al modulo di acquisizione delle liste d'autorità. Il modulo

di acquisizione delle liste di autorità restituisce la lista di autorità al modulo di interfaccia utente che, a sua volta, la restituisce al browser Web per la presentazione all'utente;

- l'utente consulta nel browser Web la lista d'autorità ed individua un elemento di suo interesse. A questo punto richiede tale elemento al sistema centrale;
- il modulo di interfaccia utente del soggetto conservatore inoltra la richiesta al modulo di gestione delle liste di autorità del sistema centrale. Il modulo di gestione delle liste di autorità restituisce l'elemento della lista di autorità al modulo di interfaccia utente che, a sua volta, le restituisce al browser per la presentazione all'utente;
- l'utente consulta nel browser Web l'elemento di suo interesse. A questo punto, utilizzando le informazioni ricevute, richiede la creazione di una nuova descrizione nell'archivio del soggetto conservatore;
- il modulo di interfaccia utente del soggetto conservatore inoltra la richiesta al modulo di gestione delle descrizioni degli archivi. Il modulo di gestione delle descrizioni degli archivi provvede a memorizzare la nuova descrizione e restituisce una conferma dell'avvenuta creazione al modulo di interfaccia utente che, a sua volta, la restituisce al browser Web per la presentazione all'utente;
- l'utente conclude l'interazione con il sistema.

Ringraziamenti

Questo lavoro è stato parzialmente supportato da una cooperazione tra l'Università degli Studi di Padova e la Regione del Veneto nel contesto del progetto "Sistema Informativo Archivistico Regionale" (SIAR). Il lavoro è stato inoltre parzialmente supportato dalla rete di eccellenza sulle digital library DE-LOS, come parte del programma *Information Society Technologies* (ISTI) nel sesto programma quadro della Commissione Europea (contratto G038-507618).

Maristella Agosti – Nicola Ferro – Gianmaria Silvello*

* Università degli Studi di Padova – Dipartimento di Ingegneria dell'informazione.