

# Ruolo e funzionalità dei depositi istituzionali

*Uno spazio virtuale al servizio  
della comunicazione scientifica*

Anna Maria Tammaro  
Università degli studi di Parma  
annamaria.tammaro@unipr.it

Teresa De Gregori  
Biblioteca di Economia,  
Università di Roma "Tor Vergata"  
de.gregori@biblio.uniroma2.it

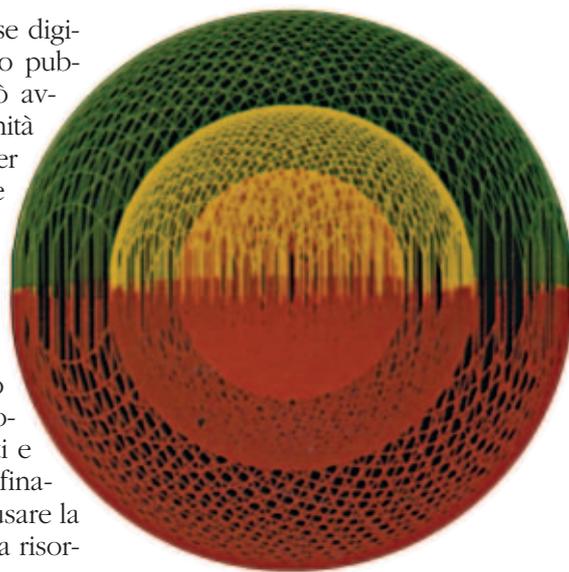
Cosa sono i depositi istituzionali? Per chiarezza espositiva, è più facile iniziare il discorso dicendo che cosa non sono. Il concetto di deposito istituzionale non è quello dell'archiviazione di una collezione di documenti su supporto digitale. Questa è infatti l'opinione più diffusa, anche tra i bibliotecari che dovrebbero essere i più esperti. Questa opinione non solo limita le potenzialità dei depositi istituzionali ma rappresenta l'ostacolo maggiore al loro sviluppo in vere e proprie biblioteche digitali, così che le risorse impiegate rischiano davvero di essere sprecate senza ottenere i benefici desiderati per i destinatari del servizio. Un altro concetto di deposito istituzionale è strettamente legato ai mezzi alternativi di diffusione della comunicazione scientifica, e questo aspetto è sicuramente corretto nel delineare una delle maggiori opportunità dei depositi istituzionali, ma fuorviante circa la loro funzione che non è certo quella di "pubblicazione editoriale". Infine i depositi istituzionali vengono di solito attivati da istituzioni scientifiche, come università e centri di ricerca, ma non sono da considerare di esclusivo interesse di tali istituzioni;

Il preprint di questo articolo, in forma provvisoria e senza la tabella di comparazione dei principali sistemi di gestione, è presente nel deposito istituzionale dell'Università di Parma curato da Anna Maria Tammaro.

anzi ogni produttore di risorse digitali o digitalizzate, soprattutto pubblico ma anche privato, può avvantaggiarsi delle opportunità dei depositi istituzionali per rendere disponibili le proprie collezioni.

I depositi istituzionali devono considerarsi vere e proprie biblioteche digitali, che si basano su un intero assetto digitale (*digital asset*): uno spazio virtuale in cui collezioni digitali, comunità di utenti e servizi interagiscono, con la finalità di creare, condividere e usare la conoscenza, rappresentata da risorse digitali. Questa definizione vuole spostare l'attenzione dall'archiviazione di una collezione ai servizi per gli utenti finali e focalizza nella condivisione di informazione la finalità di base per lo sviluppo dei depositi istituzionali.

Come le biblioteche tradizionali, i depositi istituzionali perseguono l'obiettivo di organizzare la conoscenza, soprattutto attraverso i sistemi di indicizzazione e di catalogazione. La differenza è che nei depositi istituzionali l'utente, come autore o come lettore, è parte attiva del workflow per la creazione del sistema ed è il principale gestore anche dell'accesso (si parla giustamente di "MyLibrary"). Ci limitiamo qui a dire che ogni deposito istituzionale, che spesso nasce come estensione dei servizi della biblioteca tradizionale, dovrà prevedere un'attività iniziale di indagine e coinvol-



gimento degli utenti destinatari, il cui risultato sarà una "carta" scritta (o uno statuto) in cui siano indicati la missione, gli scopi e gli obiettivi che il deposito istituzionale vuole perseguire, mentre altri documenti scritti raccoglieranno le regole comuni su cui la comunità degli utenti può accordarsi. Lo scopo delle istituzioni che realizzano un deposito è quello di utilizzare al meglio le possibilità di condivisione e collaborazione offerte dal web, fornendo un'infrastruttura informativa e comunicativa che serve agli utenti per svolgere la propria attività di studio, di ricerca, di didattica, al meglio delle possibilità. Il successo dei depositi istituzionali si dovrà misurare nel miglioramento della produttività e della qualità delle istituzioni, oltre che nell'impatto che essi avranno nel cambiamento orga-

nizzativo delle istituzioni che li rendono disponibili come servizio.

In questo articolo si intende dimostrare l'interrelazione e la complessità degli elementi in gioco, sia dal punto di vista dei requisiti tecnici che da quello organizzativo, e come il valore reale dei depositi istituzionali vada ricercato nelle enormi possibilità di condivisione delle risorse e di comunicazione interpersonale tra gruppi con interessi simili. A questo scopo verranno descritte le funzioni che un sistema di gestione di depositi istituzionali, qui definito come sistema di gestione di un assetto digitale (DAMS, Digital Asset Management System), deve realizzare.

### 1. Depositi istituzionali: come e perché?

È un momento di grande popolarità per i depositi istituzionali. Dopo l'Open Archives Initiative (OAI)<sup>1</sup> e la Budapest Open Access Initiative (BOAI),<sup>2</sup> il movimento degli archivi aperti (Open Archives) è stato riconosciuto come un modello innovativo di comunicazione scientifica, un nuovo canone bibliografico fondato sul modello Open Access, o accesso aperto. In sintesi, il modello Open Access rappresenta:

- un modello tecnologico, basato sull'OAI e sul protocollo OAI-PMH (Protocol Metadata Harvesting);
- un modello giuridico, che regola l'accesso libero attraverso le licenze Creative Commons;
- un modello economico, libero per l'utente finale e in cui paga l'autore/l'istituzione;
- un modello organizzativo, che distingue i Data provider dai Service provider e stimola alla collaborazione intra e interistituzionale.

In ambito accademico, dove sono stati sviluppati i depositi istituzionali, i complessi aspetti della comunicazione scientifica trovano per la prima volta una soluzione diversa –

o, per alcuni, alternativa – da quella tradizionalmente basata sulle pubblicazioni a stampa, che fa intravedere grandi promesse, come quella di migliorare l'impatto della comunicazione scientifica, ma presenta anche aspetti ancora da risolvere, come la credibilità che le pubblicazioni elettroniche debbono guadagnarsi.

Perché mai le istituzioni scientifiche (come università, enti di ricerca, fondazioni, ma anche musei, biblioteche, archivi) dovrebbero rendere disponibile un deposito istituzionale basato sull'accesso aperto? Le opportunità di un archivio aperto sono molte, ad esempio:

- offrire agli studiosi la possibilità di pubblicare facilmente i propri lavori di ricerca nel web, anche senza essere esperti di tecnologie;
- fornire un supporto ai docenti per la gestione del materiale didattico (*learning object*);
- favorire la comunicazione tra pari, oltre che la condivisione e il riutilizzo delle risorse digitali;
- consentire l'accesso più facile ed esteso a pubblicazioni e risorse digitali prodotte dall'istituzione;
- dare la possibilità di conservare per il lungo periodo la produzione intellettuale dell'istituzione;
- estendere le possibilità di ricerca delle risorse depositate negli archivi aperti, attraverso i motori di ricerca, ad esempio con la ricerca a testo pieno;
- dare la possibilità di integrare e gestire l'accesso a diversi supporti, come immagini, video e suoni;
- diminuire i costi della trasmissione dei documenti e nello stesso tempo soddisfare le richieste di copie, ad esempio per gli studenti;
- valorizzare la produzione scientifica dell'istituzione e migliorare il suo impatto sul pubblico, incluso quello per opere poco conosciute o fuori pubblicazione;
- presentare e diffondere in Internet risorse digitali fuori formato, come ad esempio le mappe;

- gestire i diritti di proprietà intellettuale e le licenze di accesso alle collezioni depositate e ai singoli documenti;

- riunire in modo virtuale collezioni disperse in diverse localizzazioni ma appartenenti a uno stesso filone, ad esempio lo stesso tema.

Vanno considerati tuttavia anche alcuni indubbi svantaggi o meglio dei rischi collegati alla pubblicazione di risorse digitali in depositi istituzionali:

- i costi correlati al ciclo di vita del documento digitale (non solo quelli legati alla prima fase di realizzazione del deposito ma soprattutto quelli relativi alla preservazione delle risorse digitali) e l'individuazione di chi li debba sostenere;

- il problema del rispetto degli IPR (Diritti di proprietà intellettuale), incluse la certificazione del documento digitale e le garanzie di non manipolazione;

- la necessità di assicurare un meccanismo di valutazione della produzione istituzionale, anche attraverso la recensione e la validazione di esperti, per garantire la qualità;
- il riconoscimento da parte della comunità accademica della stessa credibilità delle pubblicazioni elettroniche rispetto alle corrispondenti pubblicazioni a stampa.

Date queste opportunità e premesse, l'Open Access potrà sostituire la tradizionale pubblicazione scientifica a stampa? Il dibattito tra sostenitori e antagonisti del modello Open Access si basa attualmente su due prospettive opposte di sviluppo futuro, ed è indubbio che l'ago della bilancia sarà determinato dal supporto pubblico e dall'impatto politico che i sostenitori dell'accesso libero riusciranno ad avere. Per questo motivo, si tende a inquadrare il complesso fenomeno nell'ambito della crisi dei periodici e della successiva crisi dell'accesso ai periodici elettronici. Qui basti dire che i depositi

istituzionali sono sicuramente un supporto alle pubblicazioni scientifiche, favorendo la comunicazione tra studiosi e la valutazione di esperti, tuttavia non sostituiscono il processo editoriale.<sup>3</sup>

L'approccio con cui vogliamo descrivere i depositi istituzionali è basato sulle funzionalità che questi offrono e non sulla loro valenza per cambiare il modo in cui avviene la comunicazione scientifica. I depositi istituzionali devono infatti essere considerati vere e proprie biblioteche digitali, di cui hanno tutte le caratteristiche e tutte le opportunità. Si può dire che, mentre fino a pochi anni fa mancavano sistemi in grado di gestire le biblioteche digitali e si è quindi provveduto con sistemi di gestione dei contenuti abbastanza limitati nelle funzioni (ad esempio i CMS, Content Management Systems), molti dei sistemi oggi disponibili per realizzare depositi istituzionali hanno fatto tesoro degli sviluppi della ricerca sulle biblioteche digitali e si basano quindi su sistemi di gestione complessi, come i DAMS.

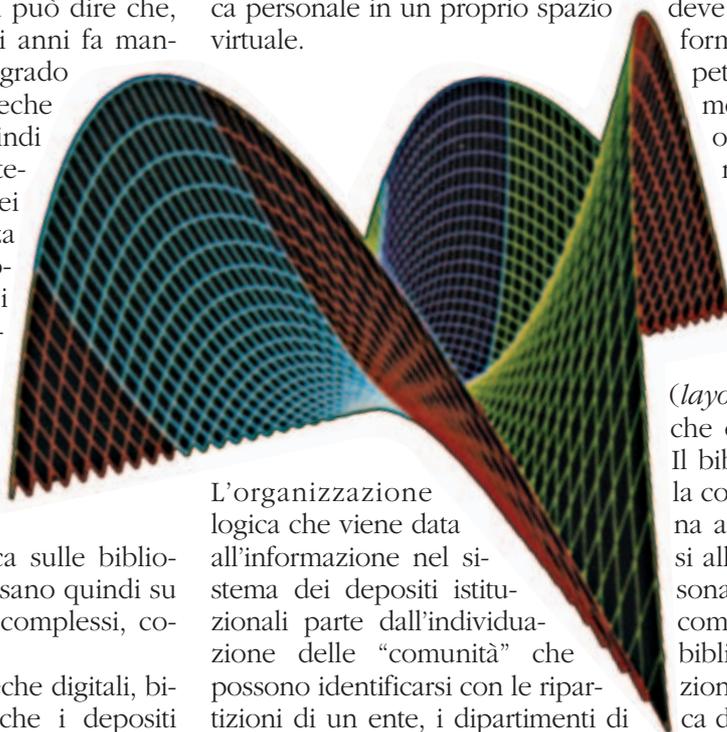
Come per le biblioteche digitali, bisogna considerare che i depositi istituzionali richiedono molto più di una semplice organizzazione di file multimediali, ma rappresentano una vera e propria organizzazione della conoscenza. Mettono quindi in primo piano una fondamentale esigenza: quella di gestire non solo le collezioni ma anche le specifiche comunità degli utenti, attori e fruitori. Questo aspetto è stato frequentemente trascurato dagli attuali sviluppatori di depositi istituzionali, e può essere considerato il punto più debole dell'attuale modello Open Access.

### 1.1. Ruoli e attori principali

I depositi istituzionali si basano su un'interazione tra gli attori in essi coinvolti. L'attore principale è sempre l'utente, che nelle biblioteche digitali è dentro il sistema, parte attiva di esso:

è *autore* del testo e dei metadata; può aggiungere delle annotazioni ai documenti; fa parte dei gruppi di valutazione e validazione dei contributi presentati;

è *lettore*, ma con la possibilità di interagire con l'autore e con gli altri utenti; può costruirsi una biblioteca personale in un proprio spazio virtuale.



L'organizzazione logica che viene data all'informazione nel sistema dei depositi istituzionali parte dall'individuazione delle "comunità" che possono identificarsi con le ripartizioni di un ente, i dipartimenti di una istituzione accademica o anche un gruppo di ricerca finalizzato allo sviluppo di un progetto. Questo approccio è significativo, in quanto nel concetto di comunità sono implicati i principi di interazione e di comunicazione, una dinamicità in cui risiede l'elemento di novità dei sistemi di gestione di biblioteca digitale di ultima generazione. Sono le comunità, attraverso i loro responsabili, che definiscono le collezioni, decidono le politiche di accesso, i criteri di validazione dei contenuti e la loro tipologia, le stra-

ategie di sviluppo. Dal punto di vista organizzativo, i depositi istituzionali rappresentano un modello di organizzazione *bottom up*, che è pragmaticamente il modo migliore per definire standard e regole pratiche per la creazione di biblioteche digitali. Dal punto di vista politico, essi rappresentano l'immagine in Internet dell'istituzione stessa, valorizzano la qualità della sua produzione, cercano di dare evidenza a quello che l'istituzione sa fare meglio.

Al bibliotecario, se esiste, o all'amministratore del sistema (che non deve essere tuttavia un tecnico informatico, ma qualcuno con competenze bibliografiche e documentarie) spetta un ruolo di coordinamento e armonizzazione, la responsabilità delle scelte in merito ai criteri e alle metodologie standard per la descrizione dei documenti, la validazione e l'integrazione dei metadata inseriti dall'autore, lo stile (*layout*) dei documenti, le politiche di conservazione dei formati. Il bibliotecario deve anche curare la comunicazione interna ed esterna all'istituzione, organizzare corsi all'utenza e curare i servizi personalizzati. Ne risulta un ruolo completamente rinnovato per la biblioteca all'interno di una istituzione, nel caso in cui la biblioteca digitale nasca e si inserisca come estensione dei servizi della biblioteca esistente, in quanto diventa struttura di supporto per gran parte dell'attività istituzionale. Nel caso di istituzioni universitarie, un ruolo specifico può essere svolto dagli amministrativi e dai segretari di dipartimento, per le interrelazioni legate alle attività di ricerca e didattica.

L'amministratore tecnico del sistema mantiene i servizi e sviluppa funzionalità aggiuntive, è responsabile della preservazione (attraverso l'attivazione di procedure di migra-

zione e/o emulazione), collabora all'architettura del sistema in ambiti federati e cooperativi, gestisce le operazioni di import/export dei metadata e delle risorse, integra e sviluppa nuove interfacce web, gestisce collegamenti con altri sistemi tipo SFX.

### 2. Funzionalità e servizi dei depositi istituzionali

Sinteticamente si può dire che le funzioni dei depositi istituzionali sono due: informare e comunicare. Il primo obiettivo dei depositi istituzionali è quello di migliorare l'accesso alle risorse digitali, sia rivolto all'interno, per conoscere le competenze e le conoscenze di studiosi e ricercatori spesso ignorate dagli organi di governo delle istituzioni culturali, sia proiettato all'esterno, per valorizzare e migliorare l'impatto della produzione scientifica istituzionale. Uno dei requisiti essenziali è quindi essere un sistema di ricerca dell'informazione efficiente, capace di identificare velocemente tutte le risorse rilevanti. Sulle funzioni di ricerca dei depositi istituzionali non è tuttavia il caso di soffermarsi a lungo, in quanto sono le stesse disponibili negli OPAC più evoluti. Quello che ci sembra importante evidenziare è che tutti i depositi istituzionali sono predisposti per essere integrati da interfacce personalizzate e soprattutto per essere aggregati, attraverso l'harvesting dei metadata, da diversi Service provider e quindi da vari Indexing service. Questa funzionalità non è ancora abbastanza visibile in Internet, per l'attuale carenza di Service provider, ma la situazione sta cambiando soprattutto per l'attività dei consorzi. L'architettura corretta in futuro dei depositi istituzionali dovrà prevedere pochi depositi distribuiti in Internet come biblioteche digitali, molti portali e servizi di indicizzazione basati sulle collezioni

contenute in tali depositi e diverse visualizzazioni delle risorse digitali per approcci personalizzati. La funzionalità dell'accesso all'informazione è quindi organizzata nei depositi istituzionali soprattutto dal punto di vista del Data provider, con strumenti di supporto alla creazione, validazione, gestione dell'accesso e preservazione delle risorse digitali.

Un aspetto potenzialmente di grande valore dei depositi istituzionali è la loro capacità di facilitare ed estendere la comunicazione, per fornire agli utenti contesti più ampi di condivisione della conoscenza.

Ancora da esplorare sono le enormi possibilità di comunicazione tra "pari" (esperti, studiosi) e di cooperazione interistituzionale offerte dai depositi istituzionali, perché gran parte di essi vengono tuttora sotto-utilizzati e sono per lo più isolati.

Quando si parla di depositi istituzionali si intende necessariamente un ambito di rete (*networking*): che senso ha allora realizzare depositi isolati, frammenti in sé perfetti ma che non comunicano? Ogni deposito istituzionale, così come ogni biblioteca digitale, dovrebbe nascere come progetto comunicativo e cooperativo. Giustamente i depositi istituzionali sono considerati nelle istituzioni di ricerca, che per prime li hanno sviluppati, alla base del metodo scientifico di condivisione, confronto e comunicazione pubblica dei risultati della ricerca, svolgendo un ruolo essenziale di facilitazione della comunicazione. A molti sembra che l'esigenza di collaborazione necessaria per ogni attività che usi Internet si possa soddisfare ottimamente utilizzando standard ormai condivisi, legati al web. Invece non è sufficiente: molte scelte politiche (criteri, regole, responsabilità) e organizzative (schemi di metadata, flussi organizzativi) devono essere stabilite a monte di ogni progetto di deposito istituzionale, trovando i partner specifici

con cui collaborare (anche all'interno della stessa istituzione di appartenenza), ed estendendo quindi le possibilità della propria collezione digitale anche ad altri utenti, esterni al gruppo di riferimento e che potranno utilizzare la collezione per scopi diversi da quelli per cui il deposito viene organizzato. Detto in altre parole, i depositi istituzionali per definizione si sviluppano in un ambito collaborativo, meglio se consortile, in cui chi aderisce all'impresa volontariamente decide di condividere risorse e servizi e di cooperare.

Le forme di cooperazione possono essere le più diverse e sono almeno in parte basate sugli stessi presupposti della cooperazione tra biblioteche tradizionali, ma con alcune nuove urgenze e con le opportunità offerte dalle tecnologie dell'informazione. Le urgenze riguardano in particolare problematiche come la preservazione di lungo periodo dei documenti digitali e la gestione dei diritti di proprietà intellettuale. Le opportunità comprendono la semplificazione dei protocolli di rete e di formato di scambio dei dati, attraverso il protocollo OAI-PMH, unitamente alla chiara distinzione di ruoli tra Service provider e Data provider che sta stimolando lo sviluppo di importanti progetti di cooperazione.

Prendendo ad esempio l'architettura modulare di un sistema come OpenDLib, in cui i vari servizi (archiviazione, ricerca, disseminazione dell'informazione, gestione del workflow) dialogano tra loro attraverso un protocollo http, si possono immaginare soluzioni diverse.

Alcune istituzioni possono decidere di consorziarsi e di riunire le risorse, anche umane, per la gestione del sistema, centralizzandolo: ad esempio attraverso consorzi inter-universitari. Ci sarà dunque un'installazione singola del sistema, completo di tutte le funzioni, un solo deposito per l'archiviazione

dei documenti, un'unica interfaccia per l'utente. Alcune funzioni possono essere replicate su altre macchine, per garantire l'accesso anche in caso di malfunzionamento di una macchina.

In altri casi, può essere invece più conveniente distribuire i moduli e i depositi su macchine diverse, dislocate presso le varie istituzioni facenti parte del consorzio, eventualmente con interfacce distinte e personalizzate. Devono tuttavia essere garantite funzionalità comuni basate su accordi collaborativi tra istituzioni e standard condivisi.

Un altro aspetto importante della cooperazione, reso possibile dall'utilizzo di sistemi open source su cui i depositi istituzionali spesso si basano, è lo sviluppo condiviso dei software e il loro arricchimento con sempre nuove funzioni. Uno sviluppo che nasce direttamente dalle esigenze emerse nell'utilizzo delle procedure della biblioteca digitale, dai bisogni espressi dagli utenti e dall'evolversi delle modalità di lavoro. Le funzionalità di base che devono essere assicurate per la gestione delle biblioteche digitali sono rese disponibili da software open source; per ulteriori funzionalità o per l'aggregazione a valore aggiunto di funzionalità diverse, i depositi istituzionali sono aperti a collaborazioni anche con ditte commerciali. In questo senso, è esemplificativo il progetto DSpace, guidato dal Massachusetts Institute of Technology in collaborazione con HP. La DSpace Federation riunisce tutte le istituzioni che hanno adottato il DSpace e si propone di condividere gli sforzi per il perfezionamento del sistema, per la sua applicabilità alle diverse piattaforme hardware e software, per lo sviluppo di nuove funzioni, oltre che per condividere la ricerca in atto sui problemi della conservazione del digitale e sull'interoperabilità fra diversi archivi. Il gruppo è particolarmente attivo, la

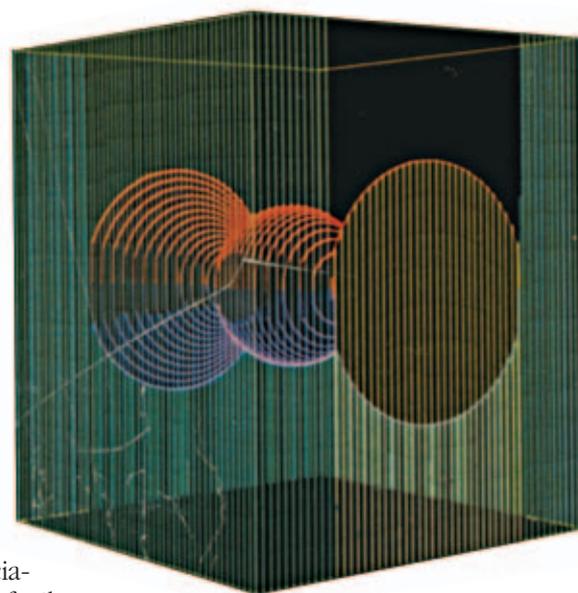
mailing list degli utenti costituisce di fatto anche un supporto tecnico rapido ed efficace per i nuovi utenti, oltre ad essere il luogo in cui vengono discussi problemi, ipotizzate nuove soluzioni ed estensioni delle funzionalità esistenti, grazie all'apporto di tutti coloro che ci lavorano.

### 2.1. Gestione del workflow

La gestione del *workflow* (o flusso di lavoro dell'oggetto digitale) è una delle funzioni essenziali dei DAMS. Facciamo un esempio concreto, di facile comprensione per tutti i bibliotecari. Nelle biblioteche tradizionali, l'organizzazione si è sviluppata attorno a una particolare catena, chiamata *iter* del libro. Il libro, come una pratica, passa da vari uffici, in cui persone con diverse responsabilità e competenze attuano il loro compito specifico. Lo scopo dell'*iter* del libro è quello di facilitare il recupero più facile e più veloce possibile dell'informazione da parte dell'utente (o se si vuole è funzionale all'*information retrieval*): quanti degli addetti in questa catena lo sanno? La frammentazione del flusso in responsabilità diverse ha infatti portato a segmentare l'*iter* in tanti step, in cui ciascuno sa solo della sua piccola parte, perdendo di vista lo scopo finale del processo.

Il *workflow* della risorsa digitale sostituisce l'*iter* del libro, ma non è una catena di montaggio. Le diverse fasi e i diversi attori responsabili lavorano in un flusso elicoidale che è la caratteristica specifica delle biblioteche digitali, in cui ognuna delle fasi previste deve necessariamente tener conto di tutte le altre. Gli step sono necessariamente integrati e prevedono:

- 1) inserimento del documento (*submission*)
- 2) peer review



- 3) metadata editing
- 4) validazione
- 5) archiviazione nel deposito
- 6) gestione identità, integrità e completezza del documento
- 7) preservazione
- 8) ricerca
- 9) navigazione
- 10) gestione delle licenze di accesso
- 11) personalizzazione
- 12) comunicazione e messaggistica
- 13) lavoro di comunità.

Esaminiamo come si articola questo percorso in DSpace. Per ciascuna collezione creata si stabilisce quali utenti hanno l'autorizzazione (*add permission*) a immettere (*submission*) i documenti per la "pubblicazione", ad esempio tutti i docenti e i ricercatori di un dipartimento. L'utente autorizzato che inizia l'immissione di un documento nella collezione deve anzitutto inserire i metadata, attraverso una maschera molto semplice che gli viene proposta, avendo a disposizione per ogni elemento un help per la compilazione; quindi passa all'invio del/i file che compongono il documento. Infine l'utente è invitato a prendere visione del documento che definisce i diritti e le licenze di accesso che vuol rendere fruibili per i contributi depositati. È

comunque possibile definire condizioni diverse per singoli documenti, su richiesta dell'autore.

Uno dei problemi emersi recentemente nei depositi istituzionali è quello di definire i criteri per la certificazione di qualità della collezione e i criteri di validazione dei singoli oggetti digitali. Inserire un documento nel web è una forma di pubblicazione e deve rispettare alcuni criteri condivisi. I depositi istituzionali devono infatti differenziarsi dai generici documenti pubblicati in rete, in quanto l'istituzione che li rende disponibili si impegna a garantire la qualità di quanto viene immesso. Tuttavia, deve essere lasciata la possibilità di immettere documenti con una post-review. Se non è stato definito un processo di peer review preventivo, il documento può essere direttamente pubblicato, cioè entra a far parte dell'archivio e viene indicizzato. Ma è possibile, e caldamente consigliato nel caso in cui l'istituzione che apre il deposito voglia mantenere la qualità della collezione, definire un workflow, considerare l'uniformità della risorsa rispetto alle politiche convenute, stabilire la conformità del documento rispetto a formati e modelli anche stilistici predefiniti, insieme alla correttezza della descrizione.

Nella prima fase della loro esistenza, molti depositi istituzionali hanno privilegiato la quantità del contenuto alla qualità, limitando l'interesse all'editing dei metadata e non utilizzando appieno le possibilità dei DAMS.

In DSpace il processo di validazione si articola in tre step, ognuno dei quali è assolutamente facoltativo. Fino a quando non è passato attraverso l'intero iter definito dal responsabile della collezione, il documento non è fruibile al pubblico, cioè non è ancora entrato nell'archivio ma risiede in una sorta di limbo, uno spazio accessibile solo

alle persone definite come attori responsabili dei diversi livelli del workflow. Solo dopo che il documento ha ricevuto l'approvazione da parte di tutti i gruppi, esso entrerà a far parte della collezione e quindi sarà inserito nell'archivio, riceverà il suo numero identificativo d'archivio e il suo identificativo permanente (*handle*), e sarà indicizzato. Il primo step di validazione previsto è quello della peer review: il documento viene valutato nel contenuto e nell'aspetto editoriale, e può essere accettato o respinto all'autore con una e-mail di notifica, contenente le motivazioni del rifiuto o i suggerimenti di modifica. Il secondo step prevede la revisione, correzione e integrazione dei metadata inseriti al momento della sottomissione del documento da parte dello stesso autore. Nessuna modifica può essere effettuata sul documento. Il terzo step, infine, prevede un ultimo controllo formale prima della definitiva accettazione. Nella seconda e terza fase il documento non può più essere respinto, ma solo rinviato allo step immediatamente precedente. Quando un nuovo documento viene proposto in una collezione, e quando passa da uno step del workflow al successivo, il sistema provvederà a inviare alle persone definite come attori del processo una e-mail che notifica la presenza di un nuovo processo in atto (*task*), cioè di un nuovo contributo da esaminare. Nel "MyWorkspace" i revisori potranno trovare l'elenco dei processi in atto, rivedere quelli già conclusi, e potranno verificare il documento proposto e accettarlo inviandolo allo step successivo, oppure respingerlo. Quando una persona del gruppo dei revisori prende visione del documento, questo viene tolto dallo spazio di attesa (*pool*) per evitare che più persone possano esaminare contemporaneamente lo stesso documento, ottimizzando così i tempi di lavoro.

Scholnet ha sviluppato una funzione che consente a un utente autorizzato di aggiungere annotazioni a un documento. Questo può avvenire in varie situazioni: durante il workflow, su un documento non ancora "pubblicato", o su un documento già parte di una collezione. L'annotazione, qualificata secondo il contenuto (di critica, di commento, per creare un legame con altri documenti ecc.), sarà accessibile, anche attraverso chiavi di ricerca, agli utenti autorizzati. Può nascere così un dibattito attorno ai contributi archiviati, che può dar luogo a nuove formulazioni e nuova conoscenza. In tale modo si realizza all'interno del sistema un lavoro di creazione intellettuale, di sviluppo della ricerca.

Queste funzioni possono essere utilizzate anche per realizzare percorsi didattici, archiviando in una collezione riservata le dispense di un corso cui gli studenti possono accedere, e attraverso annotazioni e commenti interagire con il docente. Gli studenti hanno la possibilità di produrre degli elaborati che saranno valutati dal docente, e svolgere lavori di gruppo.

### 2.1.1. Metadata

La funzione dei metadata è essenziale per la biblioteca digitale e per l'OAI. Essi non si limitano alla descrizione bibliografica ma interessano ogni aspetto gestionale e amministrativo per l'accesso alla risorsa digitale. I metadata sono quindi utili non solo per la fase di accesso ma per tutte le funzioni e i servizi che caratterizzano la biblioteca digitale. Inoltre servono anche per le collezioni, oltre che per i singoli documenti. A seconda della granularità scelta, saranno inseriti metadata anche per identificare la struttura della risorsa digitale e infine metadata saranno raccolti per gli utenti, per i gruppi, per le comunità, per gli autori.

I metadata devono essere il più possibile standard: gli innumerevoli profili di applicazione sviluppati a partire da Dublin Core costituiscono esempi di come sia possibile rappresentare la specificità di ambiti disciplinari diversi, basandosi su un modello e una struttura comune (lo standard Dublin Core), con l'aggiunta di elementi e/o qualificatori. È necessario però provvedere a implementare nel proprio sistema il protocollo OAI-PMH, per garantire l'esportabilità dei dati, costruendo dei crosswalk che hanno il compito di tradurre il set di elementi utilizzato nel set previsto dal protocollo. Lo scopo è l'interoperabilità e l'integrazione necessaria nel modello Open Access.<sup>4</sup>

I metadata per la semantica sono quelli ritenuti attualmente più importanti per l'identificazione dei contenuti delle risorse digitali; tuttavia sono quelli che danno le maggiori difficoltà di standardizzazione. Ontologie, thesauri, classificazioni sono gli strumenti attualmente più usati nelle biblioteche digitali. Una particolare attenzione andrebbe dedicata anche a quegli elementi dello schema di metadata in uso che contengono l'esplicitazione di una relazione di qualche tipo con altri documenti, per consentire di ricostruire il contesto dell'informazione, l'evoluzione di idee e progetti, per esplicitare e rendere navigabile quella fitta rete di legami con altre conoscenze, insita in ogni espressione dell'ingegno.

Da una parte, sia nel dibattito sui metadata che nello sviluppo dei nuovi sistemi di gestione delle biblioteche digitali, si nota la tendenza ad accogliere una molteplicità di dati descrittivi e gestionali; dall'altra però questi sistemi non sviluppano abbastanza funzionalità accessorie per il trattamento dei dati stessi, come ad esempio permettere di creare e mantenere liste di autorità o di implementare schemi controllati in modo dinamico, rendendo cioè

possibile navigare attraverso la rete delle relazioni semantiche che collega tra loro le varie voci dei dizionari, e arrivare così ai documenti. Parallelamente, anche le funzioni di indicizzazione sono oggi molto semplici. Ovviamente, l'architettura aperta di tali sistemi consente di sviluppare funzioni ancora non previste, ma lo stato attuale di sviluppo fa percepire tale incertezza tra la semplificazione a favore di un accorciamento dei tempi e di un risparmio sui costi per il trattamento dei documenti, e la complessità e analiticità della descrizione per fare emergere tutte le potenzialità insite nei documenti a favore di ricerche più complete ed esaustive.

Elementi di particolare rilevanza negli schemi di metadata sono quelli che individuano i diritti di proprietà intellettuale e il tipo di accesso disponibile per ciascuna risorsa digitale e per ciascuna collezione dei depositi istituzionali.

Ci sono attualmente diverse opzioni per descrivere i metadata che identificano le licenze di accesso dei documenti digitali. Tra queste la Creative Commons Initiative (2002) è specificamente nata per il modello Open Access.<sup>5</sup> L'iniziativa ha delineato una serie di undici diverse licenze di accesso, con cui gli autori possono rendere disponibile la propria opera. Queste licenze hanno tre descrizioni: una semplice *human-readable*, con delle icone di comprensione intuitiva, una *lawyer-readable* per gli avvocati e una *machine-readable rights metadata*.

I diritti di proprietà e le licenze di accesso sono divisi in cinque approcci:

- IPR sui metadata (*Default repository-wide rights expressions over metadata*);
- IPR sulle risorse (*Default repository-wide rights expressions over resources*);
- IPR opzionali di collezione (*Optional set-level rights expressions over resources*);

– IPR sui singoli record (*Rights expressions over individual metadata records*);

– IPR su singole risorse (*Rights expressions over individual resources*).

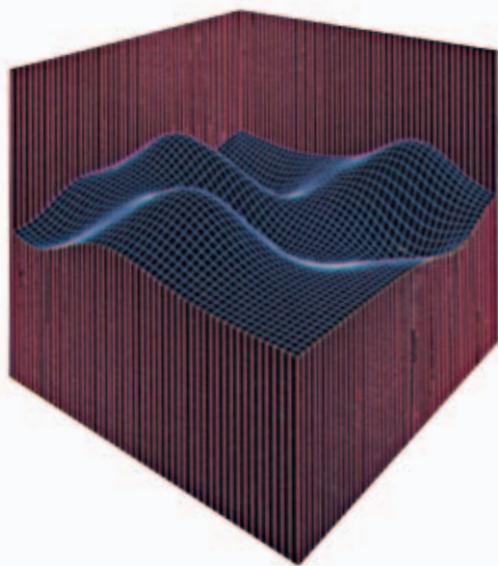
Ci sono infine i metadata amministrativo-gestionali (MAG), di cui fanno parte anche i metadata sui metadata, che hanno il compito di registrare informazioni sulla provenienza, l'integrità, i diritti sui metadata che accompagnano un documento reperito nella rete: ci sono quindi elementi che registrano dati sull'istituzione responsabile per la produzione dei metadata sul documento, il tipo di azione di cui è responsabile (creazione o solo modifica), data dell'azione ecc.

Per chiarezza sia degli utenti sia dei fornitori di servizi e informazioni è raccomandato nell'OAI di distinguere i metadata dalla risorsa. L'inserimento dei metadata è una delle funzioni in cui è importante trovare l'equilibrio tra costi e benefici, anche considerando la possibilità di condividere l'attività in ambiente di consorzi e federazioni, costruendo sulle migliori esperienze dell'automazione bibliotecaria. La funzione di editing dei metadata centralizzata, ad esempio, potrebbe essere abbinata all'inserimento di metadata distribuiti o generati automaticamente (import/export, harvesting).

## 2.2 Gestione dell'accesso e della preservazione

Le funzionalità di ricerca dei depositi istituzionali sono volutamente limitate a funzioni di ricerca e navigazione standard, tuttavia vale la pena segnalare alcuni strumenti più sofisticati.

Ad esempio in OpenDLib è possibile la *cross language search*, che consente la ricerca di un termine nelle varie lingue in cui esso può apparire all'interno dell'archivio, grazie all'utilizzo di dizionari in fase di indicizzazione e recupero; e il



*relevance feedback*, che è un metodo per calibrare la ricerca sui risultati giudicati pertinenti di una ricerca precedentemente effettuata. Ciò che acquista nuova valenza è la funzione di visualizzazione del record o *item*, come spesso viene definito l'insieme della descrizione o delle descrizioni del documento e dei file che lo compongono. L'oggetto informativo è infatti spesso molto complesso e di questa complessità deve render conto la visualizzazione completa dell'*item*, per consentire una comprensione esatta della struttura logica del documento nel suo insieme.

In OpenDLib, dalla visualizzazione breve dell'*item* si può passare alla visualizzazione completa che appare come un indice strutturato in cui compaiono le descrizioni del documento, secondo lo schema di metadata usato; le suddivisioni logiche del documento, in capitoli, parti ecc.; gli allegati; le annotazioni; il reference linking; ciascuno con le varie versioni disponibili (pdf, html, txt ecc.).

I depositi istituzionali, dal punto di vista dei Data provider, focalizzano la preservazione come funzione prioritaria per l'accesso. Il problema della preservazione non è alla fine del ciclo del documento, come succede per la biblioteca tradizio-

nale, ma è il presupposto per l'accesso alle risorse. In particolare la preservazione di lungo periodo non deve intendersi limitata alla preservazione dei supporti fisici di memorizzazione. La conservazione del digitale consiste nella pianificazione, nello stanziamento delle risorse e nell'applicazione dei metodi e delle tecnologie per la conservazione, necessari ad assicurare che l'informazione digitale di valore duraturo (*continuing value*) rimanga accessibile e utilizzabile. Un elemento di riflessione necessaria riguarda la selezione dell'informazione da conservare, quella cioè che avrà un'utilità, un senso anche in futuro.

A livello globale, il dibattito è aperto e le iniziative sono diverse, ma anche nelle istituzioni, accademiche o pubbliche, c'è dell'informazione la cui utilità futura è dubbia, e considerando gli alti costi che anche la conservazione del digitale comporta, è bene che sia trattata in modo diverso.<sup>6</sup>

DSpace, oltre alla possibilità di utilizzare tutti i metadata amministrativo-gestionali che si ritengono necessari (salvo poi, però, provvedere alle funzioni che ne permettano un reale utilizzo, dalle procedure per renderli ricercabili a eventuali funzionalità gestionali collegate), mette a disposizione uno strumento che ha lo scopo di rendere più agevole le operazioni di conservazione dei documenti, sotto il profilo del controllo dei formati: questo strumento è il registro dei formati (*bitstream registry*), in cui vengono registrati i formati per i quali l'istituzione si impegna a garantire un'operazione di conservazione perenne, con le tecniche che si riterranno più opportune anche riguardo alla tipologia dei documenti in questione (refreshing, emulazione ecc.).

L'autore che introduce un documento per l'archiviazione può prendere visione del registro, in

modo da essere consapevole del "destino" futuro del proprio lavoro, ed eventualmente presentarlo in un formato diverso, scegliendo tra quelli per i quali è garantita la conservazione perenne.

Un altro strumento atto a garantire l'accessibilità futura dei documenti è l'utilizzo degli identificatori univoci, che sono una cosa diversa da non confondere quelli attribuiti localmente dal sistema: parliamo qui di identificatori che hanno una validità globale e sono gestiti da entità che si pongono a un livello superiore rispetto ai singoli archivi.

Nel modello Open Access sostenuto dai depositi istituzionali le entità superiori, o agenzie, che attribuiscono gli identificatori sono pubbliche e in ogni caso non-profit, anche se le risorse raccolte possono avere identificatori non "open" come il DOI.

Il DSpace implementa l'*handle system*, un software che attribuisce automaticamente un numero identificativo nel momento in cui il documento è accettato e quindi viene inserito nell'archivio. Questo numero è composto da una parte fissa, che è l'identificativo attribuito all'istituzione responsabile dell'archivio dal CNRI, che gestisce il sistema in modo gratuito, e da un numero progressivo. Si tratta quindi di una sequenza numerica univoca che permane anche se la risorsa digitale viene trasferita in depositi diversi. Contemporaneamente il documento riceve una URI, che è formata dall'indirizzo IP del sistema di gestione centralizzato degli handle (hdl.handle.net), più l'handle che identifica il documento. In questo modo, l'*handle resolver* dell'agenzia non-profit CNRI sarà capace di localizzare il documento ovunque esso sia nella rete, indipendentemente da quella che è stata la sua storia (cambiamento di repository, cambiamento dell'istituzione che lo gestisce ecc.). Nelle citazioni andrà poi utilizzata sempre questa URI

per permettere anche in futuro di rendere rintracciabili i documenti citati, così come succede ora per i documenti a stampa. Vista la complessità dei problemi di preservazione, il responsabile non può essere un singolo (ad esempio un privato, o anche un'istituzione che non dia garanzie) ma un'istituzione pubblica con una prospettiva di vita permanente.

Non c'è accesso senza autorizzazione e autenticazione. Il problema non è solo tecnico, nel senso che deve essere garantito da macchine e programmi, ma soprattutto politico e giuridico.

Per questo motivo la gestione delle autorizzazioni e delle licenze di accesso è integrata nella biblioteca digitale e l'autorizzazione riguarda i singoli oggetti digitali, le eventuali collezioni, le attività consentite per ciascun utente o gruppi di utenti che devono essere riconosciuti e autenticati.

In DSpace le funzioni atte a definire politiche di accesso, privilegi e ruoli all'interno del sistema sono molto sviluppate. Tutto si gioca nel rapporto tra oggetti (i documenti), persone (*e-people*, gruppi e comunità) e azioni (lettura, scrittura ecc.), variamente combinati tra loro per dar luogo a situazioni diverse e soddisfare le varie esigenze. Le politiche di accesso possono essere definite a livello di collezione, valide dunque per tutti i documenti che ne fanno parte, ma anche a livello di singolo documento, per cui è possibile stabilire che in una collezione in cui tutti i documenti sono liberamente accessibili per l'utente non registrato (*anonymous*), alcuni siano accessibili nel testo completo solo a un gruppo di utenti autorizzati. Le attività possibili per gli utenti sono: ricerca e navigazione, lettura e downloading, manipolazione e ri-uso, personalizzazione, intesa come

uno spazio che l'utente si crea nel sistema non soltanto per memorizzare percorsi di ricerca, scorciatoie per le collezioni preferite, documenti selezionati ecc.: anche qui, l'evoluzione sta nella comunicazione. Il "MySpace" diventa anche il luogo dove comunicare con gli altri, uno spazio di lavoro, all'interno della biblioteca, dove si discute e si crea informazione.

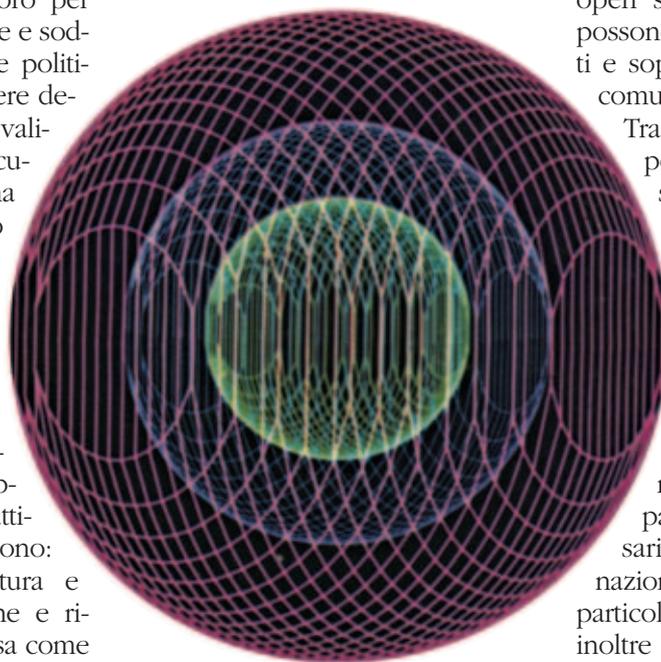
## Conclusioni

In conclusione, i depositi istituzionali rappresentano attualmente le esperienze più avanzate di biblioteche digitali, in cui il focus non è sulle collezioni (i contenuti) o sul workflow (i processi) ma sugli utenti, come comunità attive nel sistema e comunicanti tra loro (nei sistemi più evoluti). In questo senso, i depositi istituzionali hanno la potenzialità di diventare un importante strumento di cambiamento organizzativo e procedurale nelle istituzioni che rendono disponibile tale servizio. Una potenzialità non ancora pienamente attuata, perché nelle prime realizzazioni non si è tenuto in sufficiente considerazione

proprio la valenza dell'organizzazione istituzionale per il successo del deposito. Anche per la difficoltà delle biblioteche tradizionali, molto attive nello sviluppo dei primi depositi istituzionali, a inserirsi e a conoscere appieno le esigenze e le abitudini delle persone che operano nelle istituzioni, gran parte dei depositi istituzionali esistenti è povero di contenuti, oppure veicola informazioni di bassa qualità.

Per ottenere i benefici promessi dall'avvio di un deposito istituzionale bisognerà quindi basarsi sul consenso e soprattutto sul coinvolgimento delle comunità di utenti per cui si organizza il servizio. Cosa fa chi intende avviare un deposito istituzionale? Di solito parte da esigenze più o meno espresse dagli utenti. Per prima cosa si valutano i prodotti software esistenti, oppure si comincia a sviluppare un software nuovo. Questo è attualmente il problema più importante con cui ci si deve confrontare: anche chi vuole aderire a OAI, non ha chiari i ruoli di Data provider e di Service provider che il protocollo prevede e spesso non comprende tutte le funzionalità richieste per un deposito istituzionale. I programmi software open source attualmente esistenti possono dirsi completi e ben testati e soprattutto sono mantenuti da comunità autorevoli.

Tra le funzionalità che un deposito istituzionale deve consentire, quella che ci sembra essere più importante è la facilità e la spinta insieme alla collaborazione che tutti i depositi istituzionali dovrebbero prevedere come loro missione precipua. La cooperazione dovrebbe partire dall'istituzione stessa, nei suoi diversi dipartimenti e servizi, e necessariamente estendersi a livello nazionale e internazionale. Una particolare cooperazione andrebbe inoltre avviata tra depositi istituzio-



nali di università e biblioteche, archivi, musei e in genere istituzioni appartenenti ai beni culturali. Un aspetto, questo, in realtà ancora carente, forse perchè l'esperienza dei depositi istituzionali è ancora recente. C'è da augurarsi che i servizi dei Service provider ora assenti possano in tempi brevi colmare in parte la mancanza di cooperazione esistente tra depositi istituzionali, ad esempio attraverso l'attività di consorzi che si attivino per garantire una ricerca unificata, attraverso l'aggregazione dei metadata, a vantaggio della ricerca dell'utente. Molte quindi sono le opportunità dei depositi istituzionali, una tecnologia ancora da sfruttare appieno e per la quale servono volontà di apertura e di condivisione.

### Bibliografia

CERN document server software: features, CERN, 2002, <<http://cdsware.cern.ch/features/index.shtml>>.

DE ROBBIO, Antonella, *Relazione tecnica su CDSware*, 21/10/2002, <<http://eprints.rclis.org/archive/00000031/>>.

IDEM, *Relazione tecnica su DSpace (MIT)*, 23/10/2002, <<http://eprints.rclis.org/archive/00000032/>>.

*DSpace: a sustainable solution for institutional digital asset services: spanning the information asset value chain: ingest, manage, preserve, disseminate: internal reference specification: functionality*, version 01/03/2002, <<http://dspace.org/technology/functionality.pdf>>.

*DSpace: a sustainable solution for institutional digital asset services: spanning the information asset value chain: ingest, manage, preserve, disseminate: internal reference specification: technology and architecture*, version 01/03/2002, <<http://dspace.org/technology/architecture.pdf>>.

*DSpace: durable digital depository*, <<http://dspace.org/index.html>>.

*DSpace system documentation*, version 1.1-2, 03/06/2003, <<http://dspace.org/technology/system-docs/index.html>>.

GNU Eprints 2, <<http://software.eprints.org/>>.

GNU Eprints 2.2: documentation. Uni-

versity of Southampton, c 2002, <<http://software.eprints.org/documentation.php>>.

GRUPPO DI LAVORO PER L'INTEGRAZIONE DELLE RISORSE ELETTRONICHE, *Relazione finale sulla sperimentazione del sistema GNU Eprints 2*, version 1.7, 26/09/2002, <<http://strinfo.cab.unipd.it:8080/CAB/Gruppi/Documents/1043677119.41/Report.txt>>.

LYNCH, Clifford, *Institutional repositories: essential infrastructure for Scholarship in the digital age*, "ARL Bimonthly Report", 226 (2003), <<http://www.arl.org/newsltr/217/mhp.html>>.

PAGANO, Pasquale, *Open access to digital libraries: a second generation DL*, Scholnet, Pisa, ISTI-CNR, 15-16/05/ 2003 [slides].

*Scholnet: developing a digital library tested to support networked scholarly communities*, <<http://www.ercim.org/scholnet/>>; <[http://www.ercim.org/scholnet/fact\\_sheet.pdf](http://www.ercim.org/scholnet/fact_sheet.pdf)>.

TAMMARO, Anna Maria, *Publication vs communication? Searching for an identity for the University of Florence e-press*, in *Scholarly communication and academic presses: proceedings of the international conference*, University of Firenze, 22 March 2001, a cura di A.M. Tammamaro, Firenze, FUP, 2002.

VESELY, Martin, *CERN document server software* [presented at 1st OAF-workshop, 13-14 May 2002, Pisa]. <[www.oa-forum.org/otherfiles/pisa\\_cern.ppt](http://www.oa-forum.org/otherfiles/pisa_cern.ppt)>.

### Note

<sup>1</sup> L'OAI è un'iniziativa lanciata nel luglio 1999 a Santa Fé con lo scopo di garantire l'accesso tramite una comune e semplice interfaccia ad archivi di preprint, biblioteche digitali e archivi aperti di risorse digitali. L'iniziativa ha preso le mosse dagli archivi Eprint, primo tra tutti quello dei fisici di Los Alamos, ma si è estesa successivamente a diverse realizzazioni di istituzioni pubbliche e, ultimamente, è ampiamente applicata anche da privati. Un elemento essenziale dell'OAI è il Protocol Metadata Harvesting, lo schema di metadata consigliato è il Dublin Core, codificato in una speciale sintassi XML. Per ulteriori informazioni: <[www.openarchives.org](http://www.openarchives.org)>.

<sup>2</sup> BOAI è stata lanciata alla fine del 2001 sulla base della convinzione che la pro-

duzione scientifica appartenga al mondo intero e come tale debba essere liberamente accessibile al pubblico. La strategia stimolata da BOAI si fonda su: 1) auto-archiviazione dei lavori scientifici da parte degli autori; 2) periodici scientifici ad accesso libero. Ulteriori informazioni all'indirizzo: <[www.soros.org/openaccess/read.shtml](http://www.soros.org/openaccess/read.shtml)>.

<sup>3</sup> È importante distinguere tra comunicazione scientifica, che può essere basata sui depositi istituzionali, e pubblicazione scientifica, che si realizza alla fine di un processo editoriale. Le university press cercano di correggere le inefficienze dell'attuale sistema editoriale consentendo alle istituzioni accademiche di riprendere il controllo perduto delle pubblicazioni scientifiche. Per un approfondimento, si rimanda a: *Publication vs communication? Searching for an identity for the University of Florence e-press*, in *Scholarly communication and academic presses: proceedings of the international conference*, University of Firenze, 22 March 2001, a cura di A.M. Tammamaro, Firenze, FUP, 2002.

<sup>4</sup> In OpenDLib è prevista la possibilità di impiegare set differenti di metadata per diversi tipi di documenti o anche per lo stesso documento.

<sup>5</sup> La garanzia e la protezione degli IPR per l'Open Access può essere assicurata dalle licenze Creative Commons (CC). Lo sviluppo dell'espressione dei diritti CC in OAI-PMH è sostenuto da OAI/RoMEO Technical Committee, OAI-RIGHTS. Le linee guida che si stanno realizzando saranno sicuramente un incoraggiamento per gli autori e le istituzioni pubbliche a usare il modello Open Access, con i conseguenti benefici attesi per gli utenti delle biblioteche digitali.

<sup>6</sup> Anche per questo scopo esistono dei metadata. Il sistema australiano AGLS per i metadata per l'informazione governativa, ad esempio, prevede tra i MAG un elemento (*validity date range*), in cui vengono registrate delle date entro cui il documento è considerato valido. I documenti nella cui descrizione appare questo elemento saranno tenuti sotto controllo, anche mediante automatismi di sistema, per la possibile eliminazione nel tempo, o comunque per essere destinati a trattamenti di conservazione di altro tipo.

Tabella di comparazione dei principali sistemi di gestione dei depositi istituzionali

	<b>EPrints</b>	<b>DSpace</b>	<b>CDSWare</b>	<b>Scholnet</b>
Compatibilità OAI	OAI, OAIHP	OAI, OAIHP	OAI, OAIHP	Compatibilità prevista per la prossima versione.
<b>Organizzazione logico-funzionale degli oggetti nel database</b>	Si possono creare collezioni solo con l'uso di un elemento di metadata utilizzato a questo scopo.	Si possono definire "comunità" (sezioni, dipartimenti, comunità virtuali di utenti) e "collezioni", entrambe personalizzabili con diversi parametri. Sono disponibili strumenti di gestione delle collezioni.	Si configura come portale all'interno del quale sono presenti diverse collezioni, anche in database separati, strutturati diversamente.	Si possono definire collezioni di documenti. Sono disponibili strumenti di gestione delle collezioni.
<b>Gestione del documento</b>	Accoglie qualunque tipo di identificativo, ma non lo attribuisce automaticamente.	Integra l'handle system per l'attribuzione di identificativi permanenti dei documenti digitali, incluso il <i>checksum</i> . Gestisce il documento come insieme di oggetti digitali diversi, logicamente collegati.	Gestisce il sistema open URL. Timbratura digitale del documento.	Utilizza gli <i>bundles</i> .  Funzione di visualizzazione della struttura logica del documento, e delle informazioni collegate (metadata, annotazioni ecc).
<b>Gestione dei metadata</b>	È possibile definire qualsiasi schema, e per ogni tipologia di documento si può definire un set di elementi diverso, stabilendo gli elementi obbligatori e quelli opzionali.	È possibile definire qualsiasi schema; tuttavia il sistema si basa su Dublin Core e usa alcuni elementi per certe funzioni (per l'interscambio dei dati con protocollo OAI, funzioni di ordinamento ecc.).  Prevede export, import e la ricerca su schema RDF.	Il formato per l'inserimento dei dati è MARC 21; tuttavia si possono gestire database in formati diversi, definendo tabelle di conversione ai fini della ricerca e per la compatibilità OAI.  Linking dei record.  Deduzione di una citazione dai metadata che va a implementare un database in cui funziona un sistema di reference linking.  Importazione di record (e relativo full text) da vari formati, mediante mappatura e profilo di import.  Esportazione dati in XML, Unimarc.	Si può usare qualsiasi schema di metadata e ad ogni documento si possono associare descrizioni diverse, usando schemi diversi  Reference linking.
<b>Funzioni di supporto all'indicizzazione</b>	Si può definire una struttura semantica di tipo gerarchico in cui ordinare i documenti.	Si può definire una struttura semantica di tipo gerarchico in cui ordinare i documenti.	Si può definire una struttura semantica di tipo gerarchico in cui ordinare i documenti.  Consente il browsing sui soggetti, con visualizzazione espansa dei legami.	Gestione thesauri, anche più di uno, multilingue, ma non integrata, nella configurazione base.  Permette importazione batch di vocabolari.  Il thesaurus può essere collegato a parole del titolo o altri campi della descrizione, o solo al soggetto.
<b>Gestione di preservazione e trattamento dei formati dei documenti</b>	Accetta qualsiasi formato; ciascun documento può essere archiviato in più di un formato.	Accetta qualsiasi formato; ciascun documento può essere archiviato in più di un formato. Gestione di un mirror per la preservazione. Funzioni automatiche destinate alla preservazione e aggiornamento degli oggetti digitali: – attribuzione di un identificatore unico; – monitoraggio di uso; – memorizzazione della storia del documento e dei cambiamenti eseguiti; – documenta i meccanismi di autenticazione.	Accetta qualsiasi formato; ciascun documento può essere archiviato in più di un formato.  Funzioni di conversione da un formato a un altro, disponibili per l'utente registrato che sottopone un documento.	Accetta qualsiasi formato; ciascun documento può essere archiviato in più di un formato.

	<i>EPrints</i>	<i>DSpace</i>	<i>CDSWare</i>	<i>Scholnet</i>
<b>Funzioni di ricerca e browse</b>	<p>Ricerca solo sui metadata, non sul testo completo.</p> <p>Ogni campo può costituire un indice da poter visualizzare e scorrere.</p> <p>Ordinamento dei risultati.</p> <p>Memorizzazione delle ricerche effettuate per gli utenti registrati.</p>	<p>Ricerca su full text e sui metadata.</p> <p>È possibile limitare la ricerca a una o più collezioni, selezionate da una lista.</p> <p>Si può effettuare la ricerca su un insieme di risultati di una ricerca precedente.</p> <p>Memorizzazione delle ricerche effettuate per gli utenti registrati.</p> <p>Si possono usare filtri.</p> <p>Ordinamento dei risultati.</p> <p>Salvataggio o invio per e-mail dei risultati della ricerca.</p> <p>Export dei metadata per i risultati della ricerca.</p>	<p>Ricerca simultanea su archivi diversi.</p> <p>Ricerca sui metadata e su full text, in vari formati.</p> <p>Ogni campo può costituire un indice da poter visualizzare e scorrere.</p> <p>Rappresentazione complessa dei risultati della ricerca, suddivisi in collezioni e secondo ulteriori criteri definiti.</p>	<p>Ogni campo può costituire un indice da poter visualizzare e scorrere.</p> <p>Ricerca multilingue (metadata e thesauri).</p> <p>Raffinamento dei risultati della ricerca attraverso il <i>relevance feedback</i>.</p> <p>Navigazione del thesauri.</p> <p>Ordinamento dei risultati.</p> <p>Diversi formati di visualizzazione dei risultati della ricerca.</p>
<b>Immissione dei documenti</b>	<p><i>Presentazione</i></p> <p>Si può caricare un file singolo o un insieme zippato oppure indicare un URL che localizza il documento remoto.</p> <p><i>Approvazione</i></p> <p>Il documento è depositato in un'area temporanea, prima dell'approvazione</p> <p>Il processo di validazione si avvale di interfaccia web.</p> <p>Processo di approvazione in due livelli (opzionali): supervisor e editor.</p> <p>Non c'è alert per editor e supervisor, per segnalare articoli nuovi da valutare.</p>	<p><i>Presentazione</i></p> <p>L'utente può scegliere in quale collezione inserire il documento che sottopone alla valutazione.</p> <p>Facilitazioni per l'inserimento dei metadata da parte dell'autore.</p> <p><i>Approvazione</i></p> <p>Il documento è depositato in un'area temporanea, prima dell'approvazione.</p> <p>Il processo di validazione si avvale di interfaccia web.</p> <p>Ogni collezione può avere un suo processo di approvazione.</p> <p>Processo di approvazione in tre livelli (opzionali):</p> <ol style="list-style-type: none"> <li>1) <i>reviewers</i>: qualità del contenuto;</li> <li>2) <i>approvers</i>: controllo editoriale e formale;</li> <li>3) <i>metadata editors</i>.</li> </ol> <p>Alert per segnalare i documenti da processare ai responsabili delle varie funzioni.</p>	<p><i>Presentazione</i></p> <p>L'utente può scegliere in quale collezione e sezione inserire il documento che sottopone alla valutazione.</p> <p>Al momento del caricamento dell'articolo, l'autore può chiedere la conversione in un formato diverso.</p> <p>Recupero dei documenti anche da altri archivi OAI e non (<i>harvesting</i>).</p> <p><i>Approvazione</i></p> <p>Il documento è depositato in un'area temporanea, prima dell'approvazione.</p> <p>Il processo di validazione si avvale di interfaccia web.</p> <p>Processo di approvazione in tre livelli (opzionali)</p> <p>L'autore e l'utente generico possono seguire il processo di valutazione del documento presentato, nelle varie fasi, tramite l'apposita segnalazione.</p>	<p><i>Presentazione</i></p> <p>L'utente abilitato può inserire documenti, compilare i metadata e assegnare un soggetto. I dati inseriti dall'autore non possono essere cambiati dall'amministratore del sistema.</p> <p><i>Approvazione</i></p> <p>Il documento è depositato in un'area temporanea, prima dell'approvazione</p> <p>Il processo di validazione si avvale di interfaccia web.</p> <p>Articolazione flessibile del processo di valutazione.</p>
<b>Gestione utenti</b>	<p>Si possono definire dei metadata per memorizzare informazioni sugli autori dei documenti.</p> <p>Registrazione degli utenti attraverso il web, mediante compilazione di una form, o via mail.</p>	<p>Si possono definire dei metadata per memorizzare informazioni sugli utenti.</p> <p>Registrazione degli utenti attraverso il web, mediante compilazione di una form.</p> <p>Si possono definire gruppi di utenti con diverse abilitazioni.</p> <p>Si possono costituire per gruppi di utenti: mailing list, bulletin board, chat, forum, newsgroup.</p>		<p>Registrazione degli utenti attraverso il web, mediante compilazione di una form.</p> <p>Si possono definire dei metadata per memorizzare informazioni sugli utenti e sugli autori dei documenti.</p>

	<i>EPrints</i>	<i>DSpace</i>	<i>CDSWare</i>	<i>Scholnet</i>
<b>Gestione dei diritti di accesso</b>	Si possono definire limitazioni all'accesso per singolo documento.	Si possono definire limitazioni all'accesso per utente, per collezione e per singolo documento, o per diverse versioni dello stesso; l'accesso è regolato per l'utente singolo o per gruppi di utenti. È possibile gestire transazioni commerciali di vario tipo (carta di credito, prepagato, abbonamento).	Si possono definire limitazioni all'accesso per singolo documento.	Si possono definire limitazioni all'accesso sia per collezione sia per singolo documento, o per diverse versioni dello stesso; per l'utente singolo o per gruppi di utenti.
<b>Personalizzazioni e servizi avanzati all'utente</b>	Alert via e-mail, sui nuovi documenti in archivio, secondo il profilo di interesse	MyDSpace consente di creare un accesso personalizzato per l'utente registrato, anche rispetto al suo ruolo (reader, metadata editor, supervisor, collection administrator ecc).  Possibilità di gestire liste di discussione, newsletter, chat, per gruppi di utenti.	Carrello.  Alert via e-mail.  Personalizzazione del layout.  Conversione del formato del documento al momento della sottomissione.	Alert via e-mail.  Si possono associare annotazioni ai documenti, pubbliche o ristrette a gruppi definiti di utenti.
<b>Amministrazione del sistema</b>	Parte off line, parte tramite interfaccia web.	Interfaccia web. Possono essere assegnati vari ruoli nell'amministrazione del sistema.	Interfaccia web.	Off line.  Un amministratore di sistema generale; un amministratore per ciascun servizio attivato; un gestore thesauri.
<b>Informazioni tecniche</b>	GNU/Linux			
<b>Sistema operativo</b>	GNU/Linux	Linux/Unix/OS X	GPL GNU	
<b>Gestione server</b>	Apache WWW server	Apache WWW server	Apache WWW server	Apache WWW server
<b>Programmazione</b>	PERL	Java	PHP, Python, WML	C++, Java
<b>Database</b>	MySQL	MySQL	MySQL	Proprietario
<b>Motore di ricerca</b>		Lucene	Google-like	