

I depositi istituzionali tra politiche mandatarie e strategie a sostegno dell'auto-archiviazione

MARIA CASSELLA

I depositi istituzionali sono una delle componenti più recenti, innovative e creative nella complessa struttura organizzativa delle biblioteche digitali. Nonostante siano stati espressamente concepiti per consentire l'auto-archiviazione diretta da parte degli autori, uno dei principali problemi di questo tipo di archivi è quello di riuscire ad attrarre una massa critica di documentazione scientificamente rilevante. Nel presente articolo l'autrice riflette su alcune strategie utili a riempire di contenuti i depositi istituzionali, soffermandosi in primo luogo a discutere delle politiche mandatarie istituzionali e dipartimentali. Di seguito, l'articolo analizza due fattori considerati strategici per il successo di un repository, quali l'advocacy e l'adozione di politiche programmatiche relative allo sviluppo delle collezioni.

Parole chiave: Depositi istituzionali – *Open Access* – Politiche mandatarie istituzionali – Politiche mandatarie dipartimentali

Introduzione

Nel complesso e articolato sistema delle biblioteche digitali, i depositi istituzionali (*Institutional Repositories*) rappresentano una delle componenti più recenti, creative e innovative. A differenza di altri gestionali più o meno consolidati nella “quotidiana realtà” delle biblioteche digitali quali ILS, *link resolvers*, *metasearchers* o i recenti *discovery tools*, i *repositories*, infatti, non sono concepiti per gestire e armonizzare i processi biblioteconomici o valorizzare le collezioni digitali, ma sono uno strumento per supportare il deposito, la disseminazione, l'archiviazione e la conservazione a lungo termine della ricerca prodotta in formato digitale [1] e si pongono, dunque, direttamente nel flusso della comunicazione scientifica, al servizio delle comunità di studiosi.

Ciononostante, diversamente dall'esperienza degli archivi disciplinari (ArXiv, RePec, CogPrints, E-LIS, Connecting Africa), gli archivi istituzionali stentano, tranne in alcuni casi, ad essere riconosciuti come un effettivo strumento di comunicazione dei risultati della ricerca scientifica e il principale problema, ampiamente diffuso e condiviso a livello internazionale, dei curatori di questo tipo di archivi, resta quello di riempirli di contenuti. Non è un caso che, nella vasta letteratura sull'*Open Access*, le tematiche

relative ai depositi istituzionali siano tra le più dibattute, mentre sembrano lontane le riflessioni sui depositi disciplinari che godono di un successo continuo e forse insperato tra le comunità scientifiche di riferimento a partire dal 1991, l'anno in cui Paul Ginsparg lanciava pubblicamente su un *server* di Los Alamos il primo archivio disciplinare per la fisica delle alte energie: ArXiv.

Uno sguardo alla letteratura

I depositi istituzionali sono una componente decisamente giovane nell'ambito delle biblioteche digitali. I due *software Open Source* più diffusi per la gestione dei depositi, EPrints e DSpace, sono stati lanciati rispettivamente nel 2000 e nel 2002 [2].

La letteratura sull'argomento dei *repositories* comincia a essere significativa proprio a partire dal 2002, anno nel quale viene pubblicato un contributo storico sul tema: lo *SPARC Position Paper* di Raym Crow [3].

Secondo Crow, il ruolo dei depositi istituzionali è di contribuire a riformare i tradizionali canali di disseminazione della comunicazione scientifica. Crow descrive quindi le principali funzioni di un *repository*, i vantaggi in termini di prestigio e visibilità per l'istituzione che lo gestisce e definisce il contenuto di un archivio di questo tipo come:

- *Institutional;*
- *Scholarly;*
- *Cumulative and perpetual;*
- *Open and interoperable.*

Il contributo di Crow viene seguito nel breve volgere di un anno da un altro articolo ormai ampiamente citato nella letteratura sull'argomento dei *repositories*: quello di Clifford Lynch su "Portal: libraries and the academy". Lynch definisce i depositi istituzionali come «a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members» [4] mettendo in luce proprio quella caratteristica di servizio verso le comunità di ricerca che rappresenta la sfida maggiore posta in essere da questi archivi.

Una sintetica e parzialissima panoramica sui principali argomenti trattati da chi in questi anni ha studiato il mondo complesso e multiforme dei depositi istituzionali, include: lo sviluppo degli aspetti tecnologici (Allinson, et al., 2008, Phillips et al., 2007, Van de Sompel - Lagoze, 2007), la necessità di analizzare i canali e gli strumenti attraverso i quali le diverse comunità scientifiche studiano e fanno ricerca in rete (Foster - Gibbons, 2006), il marketing dei depositi (Gierveld, 2006), i numerosi studi di caso e i progetti speciali, le guide e i manuali (Barton - Waters, 2005), gli studi sullo stato dell'arte nei contesti locali (Markey et al., 2007, Weenink et al., 2008) [5].

Nelle comunità professionali di bibliotecari o informatici, presso gli enti finanziatori della ricerca e tra gli *early adopters*, tra chi fa ricerca e chi lavora nel campo dell'editoria scientifica, i depositi suscitano un interesse notevole, pongono problemi, mettono in luce la necessità di scambiare esperienze e di condividerle, di sviluppare la cooperazione, di lavorare in sinergia a livello nazionale ed internazionale.

Gli IR: la ricerca del contenuto e le politiche mandatarie

Un tema ricorrente nella letteratura relativa ai depositi istituzionali è quello della carenza dei contenuti. Infatti, quelli che dovrebbero essere i principali alimentatori e fruitori dei depositi, ossia gli studiosi, docenti e ricercatori, per fare riferimento al solo contesto universitario, mostrano ancora una grande inerzia nei confronti della pratica dell'auto-archiviazione.

In generale, un primo problema concreto per i curatori dei depositi istituzionali è quello di diffondere la consapevolezza dell'esistenza e dei vantaggi dell'archiviazione nei depositi tra coloro che non hanno mai praticato l'auto-archiviazione e sono ignari delle tematiche dell'Accesso Aperto.

Per le comunità di studiosi che, invece, già conoscono questi strumenti, gli ostacoli al loro utilizzo possono essere rappresentati dai timori di violare i diritti di *copyright*, molto spesso ceduti in maniera esclusiva dagli autori agli editori, dalla cronica mancanza di tempo da dedicare all'esercizio del deposito – che si somma alle numerose attività praticate da chi fa ricerca e didattica –, dal fastidio che i docenti mostrano nei confronti di adempimenti di tipo amministrativo e pseudo-amministrativo, dalla resistenza al cambiamento.

Per superare queste difficoltà, alcuni importanti esponenti del movimento dell'OA (Harnad, Swan, Suber) propongono l'adozione di politiche istituzionali di deposito (*institutional mandate*) più generalmente denominate "politiche mandatarie". Una politica mandataria è una politica ufficiale «per mettere a disposizione ad accesso aperto (immediato, permanente, gratuito, a testo pieno online, per tutti i potenziali utenti del mondo collegati in rete) gli articoli referati che illustrano i risultati della ricerca» [6]. Il mandato a depositare in un archivio aperto può essere deciso dalle università o dai centri di ricerca – in tal caso si parlerà di politiche mandatarie istituzionali – o da un ente finanziatore della ricerca – ad esempio Wellcome Trust, NIH, European Research Council (*funder mandate*) –, o da un singolo dipartimento (*departmental mandate*).

Secondo la Registry of Open Access Repository Material Archiving Policies (ROARMAP), a giugno 2008 sono 18 le politiche istituzionali adottate da altrettante università ed enti di ricerca [7]. In Italia l'unica *policy* istituzionale a favore dell'*Open Access* è stata quella adottata a gennaio 2008 dall'Istituto Superiore di Sanità che

prevede l'obbligo di deposito del *postprint* delle pubblicazioni del personale dell'ISS nell'archivio digitale dell'Istituto.

Il numero esiguo di politiche mandatarie istituzionali, più che essere un segnale di scarso interesse nei confronti delle problematiche dell'Accesso Aperto, mette in luce la difficoltà da parte della *leadership* universitaria di adottare pratiche che possano essere percepite come coercitive dagli studiosi. A ciò si aggiunga il fatto che l'adozione di una politica di questo tipo a favore dei *repositories* presuppone un grande lavoro preparatorio di *advocacy* ed *education* nei confronti di tutte le comunità scientifiche di riferimento, nessuna esclusa. Ne parleremo più diffusamente in seguito.

Una *policy* che recentemente ha suscitato grande interesse e scalpore nel mondo dell'Accesso Aperto, nella letteratura professionale e nella blogosfera, è stata quella adottata a febbraio 2008 dall'Harvard Faculty of Arts and Sciences.

Il documento votato dall'Harvard FAS [8] stabilisce che:

Each Faculty member grants to the President and Fellows of Harvard College permission to make available his or her scholarly articles and to exercise the copyright in those articles. In legal terms, the permission granted by each Faculty member is a **non-exclusive, irrevocable, paid-up, worldwide** license to exercise any and all rights under copyright relating to each of his or her scholarly articles, in any medium, and to authorize others to do the same, provided that the articles are not sold for a profit. The policy will apply to **all scholarly articles** written while the person is a member of the Faculty except for any articles completed before the adoption of this policy and any articles for which the Faculty member entered into an incompatible licensing or assignment agreement before the adoption of this policy. The Dean or the Dean's designate will waive application of the policy for a particular article upon written request by a Faculty member explaining the need.

La *policy* harvardiana è la prima ad essere adottata da un'università negli Stati Uniti: non si tratta di un mandato istituzionale ma il suo valore simbolico è, soprattutto nel contesto nordamericano, assoluto [9].

La risoluzione adottata dall'Harvard FAS tocca i due elementi cruciali di una politica mandataria: prescrive che la letteratura scientifica venga resa disponibile *Open Access* e che sia l'autore a concedere tale permesso; crea una "licenza automatica" che viene concessa dallo studioso/autore, prima ancora del trasferimento di *copyright* all'editore, alla Facoltà.

Quanto al primo punto, non si può non evidenziare come l'ottica della politica mandataria sia stata capovolta da Harvard a favore degli autori che assumono un atteggiamento proattivo verso il deposito, concedendo il permesso all'istituzione per la quale lavorano di archiviare i risultati della propria ricerca scientifica (*all scholarly articles*).

Nessuna coercizione può essere percepita da una formulazione di questo tipo.

Quale che sia la forma adottata dalle politiche di deposito, va rimarcato il fatto che l'adozione di una *policy* è basata, comunque, sulla collaborazione fattiva tra gli elementi portanti di un'accademia – i docenti, da una parte, e la componente amministrativa, dall'altra – ed è preceduta da un lungo lavoro preparatorio e seguita da una faticosa e attiva opera di consolidamento per fugare i dubbi e le perplessità che possono emergere tra gli studiosi e supportare il processo di auto-archiviazione [10].

Nel caso in cui non si riuscisse ad ottenere dalla *leadership* universitaria un mandato istituzionale, l'esperienza insegna che, in alternativa, è possibile cercare di adottare delle politiche dipartimentali di deposito. Queste rispondono molto bene ai differenti approcci e alle complesse tempistiche delle comunità di studiosi verso il lavoro e la ricerca in rete, all'organizzazione dipartimentale delle università, alla richiesta di autonomia dei dipartimenti nel processo decisionale e manageriale di un *repository* e, da ultimo ma non per ultimo, al fatto che docenti e ricercatori abbiano sovente un rapporto molto più stretto e diretto con il proprio dipartimento che non con l'intera università.

Secondo la ROARMAP, alla data di giugno 2008 sono in totale quattro le *policy* dipartimentali adottate a favore del deposito negli archivi aperti [11] e, tra queste, quella della Electronics and Computer Science School dell'Università di Southampton, che solo ad aprile 2008 ha adottato anche una *policy* istituzionale.

Oltre le politiche mandatarie: l'*advocacy*

Obiettivo precipuo di una politica mandataria è raggiungere il deposito del 100% della produzione scientifica di un ente di ricerca. Tuttavia, l'adozione di una *policy* istituzionale a favore del deposito è condizione importante, ma di per sé non sufficiente ad ottenere la copertura totale ad accesso aperto della ricerca scientifica di un'istituzione. Secondo Arthur Sale, l'arco temporale necessario per raggiungere un tasso di archiviazione completo è di almeno due anni [12], ma questa previsione può variare notevolmente da un'installazione all'altra e sulla base delle diverse comunità di riferimento [13].

Una *policy* istituzionale deve, quindi, essere affiancata da una serie di strategie volte a sostenerla e a renderla efficace.

Un numero considerevole di fattori influisce infatti sulle pratiche di deposito: la pianificazione economica a lungo termine, le attività di promozione e *advocacy*, una programmazione chiara circa le collezioni e le comunità di utenti, uno *staff* di persone che lavorano a supporto dell'auto-archiviazione e alla risoluzione delle problematiche collegate con il *copyright*, la personalizzazione dei servizi, il coinvolgimento degli *early adopters* e delle figure più autorevoli di un'accademia, un programma di *user education* nei confronti di studenti e dottorandi, la diffusione delle logiche del lavoro di gruppo e

lo sviluppo di sinergie con altri settori portanti di un'università: ad esempio, il settore ricerca.

Ognuno degli aspetti sopra elencati meriterebbe di essere analizzato in dettaglio. Di seguito ci soffermeremo invece su quegli aspetti che rientrano in un'ottica più marcatamente biblioteconomica e cioè quelli dell'*advocacy* ed il problema dello sviluppo delle collezioni.

L'*advocacy* può essere definita come quell'insieme di attività volte alla promozione e al supporto della pratica dell'auto-archiviazione. Scopo generale dell'*advocacy* è «to inform, dispel fears and provide the necessary support to bring about the populations of a repository or service» [14].

Per quanto possibile, è assolutamente consigliabile che la promozione di un deposito venga programmata nelle differenti fasi previste per la realizzazione di un progetto di questo tipo: nella fase di *start up* per supportare il progetto, in generale, il concetto di IR e le sue funzioni, in particolare, e nelle successive fasi di gestione, manutenzione e consolidamento del progetto.

Nelle fasi che precedono o sono immediatamente successive all'installazione del gestionale prescelto, gli amministratori di un *repository* dovranno preoccuparsi di:

- tenere presentazioni, formali e informali, nelle facoltà e/o dipartimenti per descrivere il progetto ed evidenziarne i vantaggi a breve e lungo termine;
- utilizzare tutti i canali informativi interni ed esterni per la promozione dell'archivio (*newletters*, *mailing-lists*, *forum*, blog, materiale pubblicitario di vario tipo, articoli, creazione di un sito web dedicato al progetto);
- sfruttare al massimo tutte le occasioni pubbliche di incontro (convegni, seminari, presentazioni, *workshop*) e le relazioni, sia personali sia istituzionali, per promuovere l'iniziativa.

Nelle fasi successive a quella di *start-up*, invece, si potranno adottare strategie di consolidamento quali, ad esempio, quella di inviare periodicamente agli autori le statistiche sui *download* delle proprie pubblicazioni, elaborando e diffondendo via *e-mail* i dati statistici sugli articoli più scaricati.

I principali referenti dell'*advocacy* sono naturalmente i docenti/ricercatori, ma non va sottovalutata l'importanza di indirizzare le attività di promozione del *repository* anche verso altri segmenti di *stakeholders*: gli eventuali finanziatori esterni, la *leadership* universitaria – Rettore, Direttore Amministrativo, Consiglio di Amministrazione, Senato, Presidi di Facoltà – il personale amministrativo e gli studenti; in modo particolare, la categoria dei dottorandi.

Verso finanziatori esterni e organi di governo dell'ateneo l'*advocacy* ha lo scopo precipuo di garantire la sostenibilità economica a breve e lungo termine al progetto di un *repository* [15].

Con riferimento a docenti, ricercatori e dottorandi, scopo dell'attività di *advocacy*

sarà di informare dei vantaggi del deposito in un archivio aperto, favorire o consolidare e, talvolta, supportare il processo di auto-archiviazione [16].

Per queste ultime categorie di *stakeholders* si dovrà, quindi, cercare di personalizzare le attività di promozione sulle esigenze di ciascuna comunità scientifica [17], distinguendo tra gli studiosi che già praticano l'auto-archiviazione e coloro che, invece, sono ignari delle teorie e delle pratiche dell'Accesso Aperto.

La conoscenza delle dinamiche interne all'organizzazione, delle modalità con le quali le comunità di riferimento lavorano e fanno ricerca in rete, nonché la scelta della corretta tempistica [18] sono in assoluto un prerequisito per definire delle strategie di *advocacy* di successo.

Per promuovere il deposito istituzionale verso la categoria di docenti e ricercatori, ci si potrà soffermare su alcuni concetti-chiave dei vantaggi dell'auto-archiviazione rispetto, per esempio, ai siti web personali e/o istituzionali, massicciamente utilizzati, talvolta personalmente e faticosamente gestiti, dai docenti per mettere in rete le proprie pubblicazioni. Tra i vantaggi più comunemente citati nella letteratura vi sono: la possibilità di archiviare immediatamente come *preprints* i risultati delle proprie ricerche, di massimizzarne la visibilità e, quindi, l'impatto citazionale, grazie anche al protocollo di interoperabilità, la gestione standardizzata dei CV personali e il riutilizzo immediato del materiale archiviato (*one input, many outputs*), la conservazione a lungo termine [19].

Quest'ultimo rappresenta, insieme all'interoperabilità, il vero punto di forza di un *repository*. Ciononostante, l'enorme enfasi posta sull'accesso e sulla funzione primaria di disseminazione della ricerca svolta da un *repository* [20], la mancanza di risorse adeguate e di una mentalità proattiva rispetto alle problematiche della conservazione del digitale, hanno fatto sì che quello della conservazione resti a oggi un aspetto ancora poco approfondito nella pratica delle varie implementazioni, come dimostra il numero esiguo di *repositories* che dichiarano esplicitamente di avere politiche o di adottare procedure per la conservazione a lungo termine. Fa riflettere il dato riportato da OpenDOAR, il repertorio degli archivi *Open Access*, secondo il quale solo il 14% dei depositi indicizzati adottano delle strategie di conservazione, mentre l'89% non dichiara o non definisce *policies* di questo tipo [21]. Eppure, nella percezione degli utenti accademici, docenti e ricercatori innanzitutto, proprio la conservazione continua a essere una delle attività sostanziali a carico delle biblioteche, cui oggi viene affidata la memoria digitale, così come in passato veniva affidata quella analogica [22]. Da un punto di vista tecnologico, la conservazione a lungo termine del materiale digitale passa attraverso la scelta di formati non proprietari (XML) o comunque consigliati (PDF/A) [23], la diffusione di standard (OAIS ISO 14721: 2003) e di schemi di metadati dedicati alla conservazione del digitale (PREMIS).

Problematiche di altro tipo sono, invece, quelle correlate con la sostenibilità economica di un archivio digitale, la sua vitalità, l'autenticità e l'integrità dei documenti, la

fruibilità da parte delle comunità di riferimento [24].

Nel promuovere un *repository*, infine, non va sottovalutata l'importanza di coinvolgere tutto il personale bibliotecario e amministrativo (compresi gli informatici) nelle attività di *advocacy* con particolare riferimento a coloro che si occupano di ricerca o di didattica nei singoli dipartimenti, dato il loro quotidiano contatto con docenti e studenti. Infatti, se è vero che i depositi istituzionali sono gestiti in modo pressoché esclusivo da biblioteche o sistemi bibliotecari [25], è altresì vero che uno strumento di questo tipo si inserisce «in un ambiente digitale integrato (quello di un ateneo, per l'appunto, o di un centro di ricerca) nel quale i flussi di lavoro e le risorse documentarie circolano e si intersecano a più livelli e in più contesti contemporaneamente (quello dell'*e-learning*, quello della ricerca, quello amministrativo, ecc.). La sinergia tra vari settori, tra differenti profili professionali (docenti, informatici, amministrativi, bibliotecari) con ruoli e competenze talvolta anche molto diversificate, lo studio dei flussi di lavoro e di ricerca, l'interoperabilità "politica", prima ancora che tecnica, diventano un fattore chiave per il successo di un archivio» [26].

Oltre le politiche mandatarie: le collezioni

Le collezioni di un *repository* sono il vero punto di forza di un archivio di questo tipo. Possono includere le seguenti tipologie di documenti:

- copie di articoli scientifici o capitoli di volumi *peer-reviewed* (*pre-print* e *post-print*);
- *technical reports*;
- *working papers*;
- *conference papers*;
- tesi e dissertazioni;
- *datasets*;
- materiale didattico (*teaching and learning material*);
- materiale multimediale.

Alla data di giugno 2008, su 1148 depositi indicizzati in OpenDOAR, il 60% dei *repositories* archivia copie di articoli scientifici *peer-reviewed*, il 50% anche tesi e dissertazioni, il 48% comprende nel materiale archiviato copie di articoli non referati e *working papers*, il 36% *conference papers*.

Le *policies* relative al contenuto possono essere tra loro molto differenti. Scrive Peter Suber a proposito delle collezioni di un archivio aperto: «I'm tempted to predict a continuing tension between the narrow conception of institutional repositories (to provide OA for eprints) and the broad conception of IRs (to provide OA for all kinds of digital content, from eprints to courseware, conference webcasts, student work,

digitized library collections, administrative records, and so on, with at least as much attention on preservation as access). But I have to predict that the broad conception will prevail. Universities that launch general-purpose archiving software will have active constituents urging them to take full advantage of it» [27]. L'idea di Suber è, in pratica, quella di sfruttare in pieno le potenzialità dei *software* nati per gestire i *repositories*, che consentono, di fatto, di archiviare diverse tipologie di formati e oggetti digitali [28] e di organizzarli in base alle esigenze di differenti comunità di utenti virtuali. È possibile, così, rispondere in modo più completo alle articolate necessità di chi fa ricerca in rete, dato il diversificarsi e moltiplicarsi delle fonti di riferimento: *preprint* e *postprint*, *conference papers*, presentazioni, blogs, wiki, *forum*, liste di discussione, ecc.

Anche in Italia prevalgono le *policies* che non limitano il contenuto dei depositi alle pubblicazioni scientifiche *peer-reviewed*: su 36 depositi istituzionali indicizzati a maggio 2008 da OpenDOAR, solo uno (Messanae Studiorum) è dedicato infatti a questo tipo di materiale. È comunque un dato di fatto, ampiamente documentato, che gran parte dell'attenzione degli studiosi, soprattutto in alcuni ambiti disciplinari (medicina, fisica), sia concentrata sul materiale scientifico "referato" [29]. Non è un caso che il European University Association Council abbia adottato marzo 2008 le raccomandazioni del suo Working Group sull'Open Access suggerendo alle amministrazioni delle università e alle conferenze dei Rettori di sostenere la creazione di depositi istituzionali e di politiche di deposito che prevedano l'archiviazione delle copie degli articoli scientifici immediatamente dopo la loro accettazione da parte dell'editore (*postprint*).

Le politiche relative alle collezioni devono essere il più possibile esplicite, chiare e, in linea di massima, condivise tra le varie comunità scientifiche, con riferimento sia alla tipologia dei documenti, sia alle collezioni che si deciderà di archiviare nel deposito. Per ciascuna tipologia di materiale documentario selezionata, è consigliabile adottare un *subset* di metadati. Infatti,

«sebbene i principali *harvester* siano in grado di operare la conversione in un formato *standard* dei metadati raccolti negli archivi istituzionali, per renderne possibile l'interoperabilità sintattica e semantica, è importante partire da una base comune. Diversamente, l'intervento dei gestori degli *harvester* per uniformare i metadati finirebbe con l'essere, su scala nazionale e internazionale, estremamente lungo e laborioso» [30].

La scelta di un *set* di metadati è stata proposta nel recente documento elaborato ad ottobre 2007 dal Gruppo CRUI Open Access "Linee guida per il deposito delle tesi di dottorato negli archivi aperti" che ha inserito al suo interno un'appendice con uno schema di metadati Dublin Core *qualified*, identificando gli elementi ritenuti obbligatori, quelli consigliati e quelli opzionali per garantire la piena compatibilità, degli archivi italiani di tesi di dottorato, con gli *standard* europei [31].

Un discorso a se stante sono quegli archivi aperti che fungono da anagrafi dalla ri-

cerca e contengono, quindi, i metadati delle pubblicazioni scientifiche prodotte nell'ambito di un'università con eventuale link al *full-text* del documento.

Si tratta, secondo chi scrive, di una strategia per riempire di contenuti i depositi istituzionali che sembra in qualche modo tradire la funzione primaria dei *repositories*, nonché i principi fondamentali del movimento dell'Accesso Aperto. In Italia un archivio di questo tipo è AIR dell'Università Statale di Milano <air.unimi.it>.

Un'ulteriore strategia relativa allo sviluppo delle collezioni archiviate nei *repositories* è quella che riguarda il materiale retrospettivo. Infatti, se è vero che scopo precipuo di un *repository* è quello di concentrarsi nella raccolta delle pubblicazioni correnti, meglio ancora se dei *preprint*, è anche vero che la *mission* di un deposito può variare notevolmente in base alle comunità di riferimento ed accogliere tra i suoi obiettivi primari anche la raccolta di materiale retrospettivo. Sono soprattutto gli umanisti ad avere la necessità di accedere a una massa critica di materiale di questo tipo, ma non solo [32].

I problemi che possono sorgere per gli amministratori di un deposito ad accesso aperto che decidano di archiviare in esso documentazione pregressa, sono sostanzialmente due: essere certi di non incorrere nella violazione della legge sul diritto d'autore e poter contare su tempo e risorse da dedicare alla digitalizzazione dei documenti.

Un programma di retro-digitalizzazione su ampia scala è stato condotto a termine dal progetto olandese "Cream of Science" sulla base di una legge nazionale che consentiva di retro-digitalizzare gli articoli pubblicati anteriormente al 1998. Il risultato di questa iniziativa è stato il recupero di 20.000 documenti nell'archivio aperto di Cream of Science [33].

Una tipologia di materiale che potrebbe trarre notevoli benefici da progetti di retro-digitalizzazione è quella delle tesi di dottorato, sulle quali si è aperto un ampio dibattito recentemente in Italia, grazie all'esperienza maturata da università quali Pisa o Bologna che hanno avviato dei progetti di pubblicazione *Open Access* delle tesi di dottorato [34] e la collaborazione delle due Biblioteche Nazionali Centrali che, così come previsto dalla circolare n. 1746 del Ministero Università e Ricerca del 20 luglio 2007, dovranno effettuare l'*harvesting* delle tesi di dottorato depositate negli archivi aperti [35].

Se è vero che l'*Open Access* potrà giovare dei numerosi progetti di digitalizzazione retrospettiva avviati praticamente ovunque da biblioteche di ogni tipo, in locale, a livello nazionale ed internazionale, da Google come da Microsoft [36], varrà anche la pena di porsi la domanda se tutto ciò che viene digitalizzato possa fregiarsi dell'etichetta di *Open Access* oppure no. La digitalizzazione retrospettiva di materiale che non è più sotto il vincolo del *copyright* con lo scopo precipuo, anche se un po' paradossale, di conservare, è *Open Access*? Sicuramente no. La conversione in formato digitale e la messa in rete delle annate retrospettive di riviste pubblicate dalle *university press* nel corso dei decenni, rientra sotto l'etichetta di *Open Access*? Molto probabilmente sì.

Scopi e funzioni del progetto, rilevanza scientifica e caratteristiche delle opere digitalizzate di volta in volta decreteranno se un progetto di retro-digitalizzazione può essere considerato *Open Access* o no. Come hanno scritto congiuntamente Peter Suber e Stevan Harnad in un comunicato del 29 aprile 2008 sui rispettivi blog, c'è grande ambiguità nell'utilizzo del termine *Open Access* e ci sono varie gradazioni di *Open Access* che gli autori riassumono nei due aggettivi: "forte" e "debole". Il primo si riferisce alla letteratura che rimuove sia le barriere dell'accesso sia quelle legali; il secondo, alla letteratura che è liberamente accessibile, ma non garantisce alcun diritto di riutilizzo dell'opera che non sia legato al principio del *fair use*. D'ora in avanti sarà più complesso fregiarsi dell'*Open Access*, "forte" o "debole" che sia, come etichetta pubblicitaria [37].

Le collezioni dei *repositories*: le prospettive future

La massa critica di contenuto scientifico archiviato nei *repositories* è un fattore chiave per il successo di questo tipo di archivi e rappresenta al momento anche il principale parametro di valutazione di un archivio aperto [38]. Di questo sono consapevoli i curatori dei depositi istituzionali, così come appare chiaro, agli esperti in materia, che il vero valore aggiunto di un deposito istituzionale è rappresentato dalla possibilità di accedere al *full-text* dei documenti. In assenza di politiche mandatarie il tasso di archiviazione delle pubblicazioni scientifiche prodotte in un'università o in un centro di ricerca all'estero si aggira intorno al 15-20% secondo la letteratura; in Italia risulta ancora più basso. Laddove le politiche mandatarie sono state già adottate è necessario, comunque, un arco temporale di almeno due anni per ottenere la copertura totale ad accesso aperto della ricerca scientifica prodotta da un'istituzione.

Tuttavia, anche se molto lentamente, i depositi istituzionali stanno crescendo sia quantitativamente sia qualitativamente. Molte giovani menti sono impegnate in tutto il mondo nello sviluppo delle funzionalità di questo tipo di archivi, il Web 2.0 sta svelando nuove strategie per avvicinare gli utenti ai servizi delle biblioteche digitali, la tecnologia, l'evoluzione dei protocolli – non solo né unicamente il protocollo di interoperabilità OAI-PMH [Open Archives Initiative - Protocol for Metadata Harvesting] [39] – e la diffusione degli *standard* stanno rendendo gli ambienti digitali sempre più integrati e interagenti, lo scambio dei documenti più semplice e meno costoso. Nel contempo, la comunicazione scientifica sta cambiando radicalmente e molto più velocemente di quanto fosse ragionevole aspettarsi: il concetto di periodico o di volume nel senso tradizionale del termine comincia a destrutturarsi, si accresce il carattere interdisciplinare della ricerca, cresce l'esigenza per gli studiosi di potere effettuare operazioni di *text-mining* e/o di *data-mining*, cresce l'interesse di alcune discipline

verso la possibilità di accedere, ricercare e conservare i dati primari della ricerca scientifica (*raw data*) [40]. Nel breve volgere di alcuni anni, le collezioni dei *repositories* avranno probabilmente una fisionomia differente. La tecnologia e le esigenze degli utenti stanno sfumando la distinzione netta tra le due strategie consigliate per il raggiungimento degli obiettivi dell'*Open Access*: l'auto-archiviazione negli archivi Aperti (*Green Road*) e la pubblicazione in riviste *Open Access* (*Gold Road*). In un imminente futuro, alle funzioni di "disseminazione", "registrazione" e "archiviazione", "tradizionalmente" assegnate agli archivi aperti, potrebbe aggiungersi anche quella di "certificazione", ultimo baluardo dell'editoria commerciale. Questo almeno sembra essere il percorso più logico e naturale per lo sviluppo strategico dei depositi istituzionali così come previsto anche da due acuti osservatori e studiosi del mondo dell'*Open Access* quali Alma Swan e Leslie Carr: «The logical strategy for a university is to use its digital research repository as the submission tool for works that the press intends to publish formally that is to provide peer-review, an imprint and copy editing services to the author». [41]

Vanno già in questa direzione il recente progetto finanziato dal JISC denominato Repository Interface for Overlaid Journals Archives (RIOJA), nonché le innovative strategie editoriali di alcune case editrici universitarie, quali ad esempio la University of California Press che nel 2007 ha lanciato il progetto *eScholarship Repository*, una piattaforma nella quale vengono archiviati *postprint*, ma anche pubblicate edizioni monografiche e riviste *peer-reviewed* ad accesso aperto. [42].

Grazie ai *repositories*, sta maturando un nuovo sistema della conoscenza, più creativo, innovativo, democratico e aperto.

Note

Tutti i siti Web sono stati consultati in data 1 giugno 2008.

- [1] È quello che sostanzialmente si legge sul sito di DSpace nella sezione "About Dspace": «DSpace **captures** your data in any format - in text, video, audio, and data. It **distributes** it over the web. It **indexes** your work, so users can search and retrieve your items. It **preserves** your digital work over the long term». <www.dspace.org/index.php?option=com_content&task=blogcategory&cid=44&Itemid=156>.
- [2] Tra gli altri gestionali utilizzati per gli IRs segnaliamo: FEDORA, CDSInvenio, OPUS. Numerosi sono anche i prodotti sviluppati in locale. Sul mercato sono inoltre disponibili alcuni prodotti commerciali. Anche in Italia gli applicativi software più utilizzati sono DSpace ed EPrints.
- [3] Raym Crow, *The case for institutional repositories: a SPARC position paper*. Washington

- D.C.: SPARC, August 2002 <www.arl.org/sparc/bm~doc/ir_final_release_102.pdf>.
- [4] Clifford A. Lynch, *Institutional repositories: essential infrastructure for scholarship in the digital age*. "Portal: libraries and the academy", 3 (2003), n. 2 p. 327.
- [5] Tra i contribute più recenti si possono citare: Julie Allinson, Sebastien François, Stuart Lewis, *SWORD, Simple Web-service Offering Repository Deposit*. "ARIADNE", 54 (2008), January <www.ariadne.ac.uk/issue54/allinson-et-al>; Scott Phillips et al., *Manakin: a new face for DSpace*. "D-Lib Magazine", 13 (2007), 11/12 <www.dlib.org/dlib/november07/phillips/11phillips.html>; Herbert Van de Sompel, Carl Lagoze, *Interoperability for the discovery use and re-use of units of scholarly communication*. "CTWatch Quarterly", 3 (2007) n. 3 <www.ctwatch.org/quarterly/articles/2007/08/interoperability-for-the-discovery-use-and-re-use-of-units-of-scholarly-communication>; Nancy Fried Foster, Susan Gibbons, *Understanding faculty to improve content recruitment for institutional repositories*. "D-Lib Magazine", 11 (2005) n. 1, <www.dlib.org/dlib/january05/foster/01foster.html>; Heleen Gierveld, *Considering a marketing and communications approach for an institutional repository*. "ARIADNE", 49 (2006), October <www.ariadne.ac.uk/issue49/gierveld>; Mary R. Barton, Margaret M. Waters, *Creating an institutional repository: LEADIRS workbook*. MIT Libraries, 2004-2005 <www.dspace.org/implement/leadirs.pdf>; Karen Markey et al., *Census of Institutional Repositories in the United States MIRACLE Project Research Findings*. February, 2007 <www.clir.org/pubs/abstract/pub140abst.html>; *A DRIVERS guide to european repositories*. Edited by Kasia Weenink, Leo Waaijers, Karen van Godtsenhoven, Amsterdam: Amsterdam University Press, 2008 <www.surffoundation.nl/smartsite.dws?ch=ENG&id=13526>.
- [6] La definizione di politica mandataria è tratta dalla traduzione curata da Susanna Mornati della pagina di presentazione della Registry of Open Access Repository Material Archiving Policies (ROARMAP), il repertorio che registra le politiche di accesso aperto ufficialmente implementate dalle istituzioni in Europa, Stati Uniti, Canada, Australia <www.aepic.it/docs/OA/dichiarazione_it.html>.
- [7] A queste però di affiancano 22 *funder mandate* e 4 politiche di deposito dipartimentali.
- [8] Disponibile al link <www.fas.harvard.edu/~secfas/February_2008_Agenda.pdf>.
- [9] Sulla scia di quanto deciso dall'Harvard FAS pochi mesi più tardi anche la facoltà di legge ha votato la stessa risoluzione. A marzo 2008 anche l'Università della California ha rimesso ai suoi organi decisionali la proposta di adottare un policy di deposito.
- [10] Ad esempio Harvard creerà un nuovo ufficio di Scholarly Communication "to support faculty, to encourage deposit, and to develop the openaccess repository." In *Open doors and open minds: a SPARC/Science Commons White Paper*. April 2008 <www.arl.org/sparc/publications/papers/opendoors_v1.shtml>.
- [11] La University of Tasmania School of Computing, La Brunel University School of Information Systems Computing and Mathematics, la University of Southampton Electronics and Computer Science School e il Laboratoire de psychologie et neurosciences cognitives.

- [12] Cfr. Arthur Sale, *The patchwork mandate*. "D.Lib Magazine", 13 (2007), n. ½ <www.dlib.org/dlib/january07/sale/01sale.html>.
- [13] Ad esempio presso l'Università del Minho: «some research centres deposit 15 to 20% of their output, others like biology, bio-engineering and civil engineering deposit between 80 to 90%. This is despite the fact that 75% of all academic staff contributed to the institutional repository in 2005» in *A DRIVERS guide to European repositories*, cit., 2008, p. 71.
- [14] Alcuni studi di caso appaiono in questo caso emblematici. Ad esempio il caso dell'Università portoghese del Minho, che è stata una delle prime università in Europa ad adottare un mandato istituzionale, fortemente voluto nel 2005 dal Rettore stesso. Questo mandato è tuttavia supportato da una serie di strategie complementari come ad esempio quella degli incentivi economici. Cfr. Miguel Ferreira et al., *Carrots and sticks: some ideas on how to create a successful institutional repository*. "D-Lib Magazine", 14 (2008), n. 1/2 <www.dlib.org/dlib/january08/ferreira/01ferreira.html>.
- [15] Si parla in questo caso di attività di promozione *top-down*.
- [16] Si parla in questo caso di attività di promozione *bottom-up*.
- [17] Fondamentale è a proposito il contributo di amministrativi, bibliotecari e informatici che lavorano presso i dipartimenti.
- [18] «Connecting Africa also pointed out the importance of choosing the right time to advocate an institutional repository or service. CA feared losing its researchers were it to involve them prematurely without the necessary infrastructure and supporting place». In *A DRIVERS guide to european repositories*, cit., 2008, p. 85. Per la scelta dei tempi è anche importante tener conto del calendario accademico e delle sue scadenze.
- [19] Il concetto di conservazione a lungo termine fa riferimento ad «un arco cronologico sufficientemente ampio da essere coinvolto nei fenomeni di obsolescenza tecnologica e di evoluzione della comunità di utenti». Giovanni Michetti, *Il modello OAIS*, in *OAIS*, a cura di Giovanni Michetti. Roma: ICCU, 2007 p. XIV
- [20] Secondo Roosendaal e Geurts (1997) la funzione di "disseminazione", dagli autori definita "consapevolezza", insieme a quelle di "registrazione" e "archiviazione" costituisce di fatto una delle tre le funzioni chiave di un deposito. Hans E. Roosendaal, Peter A. Th. M. Geurts, *Forces and functions in scientific communications: an analysis of their interplay*. 1997 <www.physik.uni-oldenburg.de/conferences/crisp97/roosendaal.html>.
- [21] Confronta a proposito il grafico di OpenDOAR *Recorded preservation policies* alla pagina <www.opendoar.org/find.php?format=charts>.
- [22] A differenza della conservazione del cartaceo, però, nel caso del formato digitale gli attori coinvolti nella conservazione sono molteplici: biblioteche, editori, enti finanziatori, archivi e gli stessi autori.
- [23] Il formato PDF/A anche se proprietario è stato recentemente adottato dalla Digital Preservation Coalition come formato utile alla conservazione dei documenti digitali.
- [24] Specifico sulla certificazione dei *repositories* e, quindi, comprensivo anche dell'analisi del

rischio relativo alla conservazione del digitale è il *toolkit* DRAMBORA (Digital Repository Audit Method Based on Risk Assessment) <www.repositoryaudit.eu>. L'*output* che si ottiene utilizzando DRAMBORA è un registro dei rischi specifici per il *repository* valutato. In Italia DRAMBORA è stato adottato da Magazzini Digitali per la certificazione di "affidabilità". Una directory internazionale sull'argomento della *digital preservation* è il portale Preserving Access to Digital Information (PADI) Australia. In Europa è attivo da alcuni anni il programma Digital Preservation Europe <www.digitalpreservationeurope.eu>.

- [25] Lo studio portato a termine presso l'Università del Michigan sui depositi istituzionali attivati o in via di attivazione negli Stati Uniti denominato Miracle (Making Institutional Repositories a Collaborative Learning Environment) ha dimostrato che la grandissima maggioranza degli archivi di questo tipo viene gestita dalle biblioteche, o dai sistemi bibliotecari, delle università, che i responsabili della loro gestione sono nel 74% dei casi i direttori di biblioteca e che una gran parte del personale degli *staff* che lavorano attivamente sui depositi istituzionali ha un profilo professionale da bibliotecario (*metadata boys*, bibliotecari di *reference*, bibliotecari addetti allo sviluppo delle collezioni, bibliotecari coinvolti nelle attività di formazione agli utenti ecc.). La media delle persone coinvolte nelle attività di gestione dei depositi è di 7,2 unità. Cfr. Karen Markey et al., cit., 2007.
- [26] Maria Cassella, *Il ruolo del bibliotecario nei depositi istituzionali*. "Biblioteche oggi", 24 (2006), n. 5 p. 5.
- [27] Peter Suber, *Open Access in 2006*. "SPARC Open Access Newsletter" 105 (2007), 2 January <www.earlham.edu/~peters/fos/newsletter/01-02-07.htm#2006>.
- [28] In effetti, chi scrive è dell'opinione che i depositi istituzionali siano un ottimo strumento di disseminazione per tutta quella letteratura grigia che fino ad oggi ha avuto una scarsissima diffusione attraverso i tradizionali canali della comunicazione scientifica.
- [29] Non è qui il caso di entrare nel merito della discussione sul valore e l'importanza del *peer-review* nell'ambito della comunicazione scientifica. Da più parti e da molti studiosi e sostenitori dell'Accesso Aperto, Harnad compreso, il *peer-review* viene giudicato come un elemento irrinunciabile. Uno studio del Primary Research Consortium pubblicato a gennaio 2008 ha evidenziato che il 90% degli studiosi considera il *peer-review* come un fattore importante che accresce la qualità della ricerca scientifica. Cfr. Mark Ware, Mike Monkam, Peer Review in Scholarly Journals - perspective of the scholarly community: an international study. January 2008.
- [30] CRUI, *Linee guida per il deposito delle tesi di dottorato negli archivi aperti*. Ottobre 2007, p. 11 disponibile all'indirizzo <www.cru.it/HomePage.aspx?ref=1149>.
- [31] Confronta su questo argomento: Paola Galimberti, *Le linee guida per l'accesso aperto alle tesi di dottorato*, in *Cristalli di esperienza: nuove prospettive e scenari per le tesi di dottorato: conservazione, accessibilità, certificazione, formati, integrazione con Open Access, giornata di studio del CNBA, Torino, 13 maggio 2008, Politecnico di Torino*, a cura di Serena Sangiorgi, p. 9-20.

- [32] Sull'argomento alcuni spunti di riflessione vengono dall'articolo di Maria Rosaria Bacchini, Gloria Castagnolo, *Le raccolte retrospettive digitali come valore aggiunto*. "Biblioteche oggi", 26 (2008), n. 4 p. 39.42.
- [33] "Cream of science" nasce nel 2005 come progetto nazionale finanziato da SURF in collaborazione con DAREnet, raccoglie e rende accessibili le pubblicazioni scientifiche di più di 200 studiosi olandesi. Il servizio è attualmente gestito dalla Royal Netherlands Academy of Arts and Sciences <www.creamofscience.org>.
- [34] In realtà l'Università di Pisa mette in rete nel suo archivio "Sistema EDT" <etd.adm.unipi.it> sia tesi di laurea che tesi di dottorato a partire dal 2000.
- [35] Un primo test di *harvesting* delle tesi di dottorato è stato effettuato dalla Biblioteca Nazionale Centrale di Firenze in collaborazione con l'Università di Bologna nei primi mesi del 2008.
- [36] A maggio 2008 Microsoft ha comunque annunciato la chiusura dei suoi progetti Live Search Books e Live Search Academic iniziati nel 2006 in competizione con Google Book Search e Google Scholar.
- [37] Confronta il comunicato sui blog di Peter Suber e Stevan Harnad rispettivamente alle pagine <www.earlham.edu/~peters/fos/2008/04/strong-and-weak-oa.html> e <openaccess.eprints.org/index.php?/archives/399-Open-Access-Strong-and-Weak.html>.
- [38] Di recente "Webometrics", un'iniziativa del Consejo Superior de Investigaciones Científicas (CSIC), il corrispondente in Spagna del CNR italiano, ha classificato gli archivi aperti sulla base del numero di *items* archiviati. Cfr. Webometrics Ranking of World Universities <www.webometrics.info>.
- [39] Tuttavia la valutazione di un archivio aperto dovrebbe prendere in considerazione anche altri indicatori di prestazione quali: gli scopi del *repository* e la sua capacità di incontrare le esigenze di una parte anche piccola di studiosi accademici; la qualità dei metadati descrittivi e amministrativo-gestionali, la sua effettiva aderenza al protocollo di interoperabilità, i servizi a valore aggiunto offerti ai docenti e ricercatori, la sostenibilità economica a lungo termine.
- [40] Mi riferisco ai nuovi protocolli in fase di sperimentazione SWORD e OAI-ORE. Il 2 Giugno 2008 è stata rilasciata la versione beta dell'Object Reuse and Exchange (OAI-ORE).
- [41] Secondo Peter Murray-Rust l'accesso ai *raw data* è tanto più importante in quanto «the most discovery takes place along the long tail of science, among small groups of laboratory scientists who need access to current data». In Carol Minton Morris, *Strands of global web knowledge come together at the third international Open Repositories Conference 2008*. "D-lib Magazine", 14 (2008), n.5/6 <www.dlib.org/dlib/may08/morris/05morris.html>.
- [42] Alma Swan, Leslie Carr, *Institutions, their repositories and the web*. "Serials review", 34 (2008) n. 1, disponibile come *preprint* nel *repository* dell'Università di Southampton <eprints.ecs.soton.ac.uk/14965>.
- [43] <www.ucl.ac.uk/lis/rioja> <repositories.cdlib.org/escholarship>.