

***Comunicazione scientifica ed editoria elettronica:
la parola agli Autori***

*L'Utente-Autore nel circuito della comunicazione scientifica:
editoria elettronica e valutazione della ricerca*

Università degli Studi di Milano - Aula Magna

20 maggio 2003

Auto-archiviazione per la ricerca: problemi aperti e sviluppi futuri

di Antonella De Robbio

***"Tra gli scienziati e il caos
non c'è che un muro di archivi, di etichette,
di libri, di protocolli, di figure e di fogli"
[Latour B. e Woolgar S., Laboratory Life:
The Social Construction of Scientific Facts.
Sage, London 1979]***

L'auto-archiviazione o *self-archiving* è un processo essenziale nel nuovo modello di comunicazione scientifica, è il primo passo entro circuito di disseminazione informativa caratterizzato dal deposito delle proprie produzioni di ricerca entro spazi "open access".

Il concetto di *self-archiving* non va confuso quello di *self-publishing* o auto-pubblicazione, intesa come *vanity press*, termine coniato da Stevan Harnad. Stevan Harnad individua "una linea di demarcazione profonda ed essenziale nella galassia Post-Gutenberg"¹ tra le opere che rientrano nel circuito commerciale in tutte le fasi che coinvolgono l'autore nella cessione dei relativi diritti dietro compenso (royalty) ed essenzialmente accessibili a pagamento (libri, giornali, programmi, musica) e le opere che rientrano nel circuito della comunicazione scientifica, il cui esempio più rappresentativo è quello degli articoli di ricerca pubblicati nelle riviste scientifiche.

I ricercatori, contrariamente a tutti gli altri autori devono i loro guadagni non alla vendita dei loro articoli scientifici, ma all'impatto dei loro articoli sulla comunità dei ricercatori, ossia al fatto di essere letti, citati e utilizzati da altri ricercatori. Ne consegue che tutti gli ostacoli all'accesso costituiti da barriere a pagamento sono ostacoli ad una crescita culturale collettiva, ostacoli alla ricerca e ai ricercatori, che sfociano in un impatto negativo in termini di benessere economico e sociale, per il semplice motivo che vanno a ledere il processo "formativo" degli individui.

Poiché le produzioni di ricerca auto-archivate massimizzano e accelerano l'impatto della ricerca massimizzando perciò l'accesso alla ricerca stessa, molti sono i ricercatori che da alcuni anni

¹ Harnad, Stevan (1990) Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry. Psychological Science 1: 342 - 343 (reprinted in Current Contents 45: 9-13, November 11 1991).

<http://cogprints.soton.ac.uk/documents/disk0/00/00/15/81/>

archiviano i loro lavori sui server delle loro istituzioni o in server disciplinari. Un noto articolo di Lawrence² apparso su Nature rileva una media del 336% (media) in più di citazioni ad articoli online rispetto agli stessi articoli pubblicati a stampa.

Vi sono a livello internazionale notevoli esperienze in ambito scientifico, in particolare per la fisica, l'economia, la matematica, la biomedicina, l'astronomia, la chimica ...

Gli articoli auto-archiviati dagli stessi autori negli archivi elettronici istituzionali possono essere accessibili a tutti, dovunque, senza che ci sia bisogno di conoscere la loro localizzazione effettiva.

Negli Stati Uniti, nell'ultimo decennio sono sorti numerosi OpenArchive di varia natura e/o tipologia, in Europa vi sono alcune esperienze notevoli e anche in Italia stanno sorgendo i primi agglomerati per certi settori disciplinari o basati su server istituzionali.

I pre-print e i lavori intellettuali considerati tradizionalmente come letteratura grigia, rapporti tecnici, relazioni ai convegni, atti di congressi, documenti progettuali, documenti pre e post pubblicazione, sono una fonte preziosa per lo sviluppo di qualsiasi settore disciplinare. Per i settori scientifici la disponibilità immediata della letteratura prodotta è ormai una necessità non più prorogabile.

Un pre-print è una tipologia di documento, distribuito in modo più o meno limitato, relativa ad un lavoro tecnico spesso in forma preliminare, precedente la sua pubblicazione in un periodico. Molto spesso però questa tipologia di materiale non rientra per nulla nei canali della distribuzione a stampa e rimane così "nascosta" e priva di un effettivo impatto nella comunità dei parlanti.

Un post-print è la versione modificata, che ha passato il comitato editoriale, che è già stata sottoposta a refereeing. Per e-print invece si intende attualmente una evoluzione del materiale pre-print, post-print, post post-print. nella sua forma elettronica. Gli e-prints sono copie elettroniche di paper accademici. Solitamente un pre-print evolve entro un articolo di periodico o entro un volume o entro un'altra forma di pubblicazione a stampa o elettronica e diviene un post-print. Tale evoluzione, soprattutto nelle pubblicazioni a stampa, comporta che il paper diviene difficilmente accessibile a causa dei prezzi sempre più alti degli abbonamenti alle riviste, senza contare i ritardi nei tempi di pubblicazione.

E' innegabile che per settori come la fisica e la matematica, l'astronomia, il pre-print sia da decenni una forma di scambio quotidiano: le comunità lavorano sui pre-print e vengono informate immediatamente circa le nuove scoperte, ne discutono, dibattono i punti pochi chiari, collaborano tra campi affini e correlati, ...la rivista è solo lo strumento formale per una comunicazione "allargata" e probabilmente in certi campi è utilizzata solo a scopo curricolare o "conservativo", quasi come un archivio che registra la tappa finale di un processo complesso che ha coinvolto tutta la comunità.

Nell'ambito umanistico, non esiste nulla di tutto ciò, non vi è nemmeno l'uso della peer-review e dei processi di refereeing in quanto le monografie sono gli oggetti preminenti della ricerca umanistica. Un Open Archive di ambito umanistico potrebbe però riguardare aspetti di ricerca particolari e l'implementazione di archivi in questo settore potrebbe essere una frontiera assai promettente.

Accanto agli e-print server disciplinari, stanno sorgendo, a livello internazionale, e-print server istituzionali, politicamente più forti, rispetto alle aggregazioni libere messe in atto, in questi anni, da comunità disciplinari; essi infatti richiedono un'organizzazione stabile entro un quadro politico di responsabilità ben definita.

I server istituzionali, di fatto, si contrappongono all'attuale modello distorto del circuito della comunicazione scientifica, imponendosi come luoghi per il semplice deposito dei materiali.

² Lawrence, S. (2001) *Free online availability substantially increases a paper's impact*. Nature 411 (6837): 521.
<http://www.neci.nec.com/~lawrence/papers/online-nature01/>

Aprire un Open Archive è compito degli atenei, degli enti di ricerca, delle istituzioni, in quanto la ricerca si svolge, si sviluppa, ma soprattutto si produce entro questi luoghi.

Vediamo cosa sono gli Open Archives, o E-prints server, i cosiddetti Data Provider. Un deposito entro l'architettura OAI Open Archive Initiative è noto come "data provider". Sono Archivi Aperti preposti al deposito dei documenti scientifici, in forma elettronica (e-prints), funzionali alla loro gestione e conservazione, muniti di metadata in formato standard ai fini dell'interoperabilità. Si dicono OAI (Open Archive Initiative) compatibili quando rispondono ai requisiti tecnici e agli standard di interoperabilità. I documenti elettronici possono essere depositati direttamente dagli autori attraverso un semplice processo noto come auto-archiviazione o da terze persone (bibliotecari). Ciascun documento depositato in un DP Data Provider è corredato da metadata in formato standard che vengono esposti per la raccolta (harvesting) da parte dei SP Service Provider "fornitori di servizi"

Gli archivi sono detti "aperti" in quanto possono essere interrogati da un servizio ("service provider") che si occupa di indicizzare i metadata raccolti dai vari archivi attraverso la funzione di harvesting. Il colloquio e lo scambio informativo di metadata è possibile solo tra archivio OAI compatibili o interoperabili.

I modelli tecnici di Open Archives sono due: modello accentrato (arXiv.org) unico grande server a carattere nazionale o internazionale, modello distribuito (RePEc) piccoli server locali sparsi. Spesso i piccoli server non sono integrati entro una rete a modello distribuito - come avviene nel caso del modello *RePEc* - e per questa ragione molti piccoli server sono stati, nel corso degli ultimi cinque anni, assorbiti da server più grandi (per esempio da *arXiv*).

A livello organizzativo le scelte possono essere:

- **Open Archive istituzionali:** l'archivio raccoglie, quale testimonianza della produzione intellettuale dell'ente, tutti i lavori di un particolare ente (università, ente di ricerca, dipartimento) o una parte selettiva dei lavori che l'ente ritiene di conservare nel deposito. In questo caso i materiali raccolti coinvolgono varie discipline.
- **Open Archive disciplinari:** l'archivio raccoglie i lavori in una determinata disciplina. Può anche trattarsi di un server di un ente che decide di aprire più archivi per discipline differenti. Molto spesso però si tratta di più soggetti (enti o anche soggetti individuali, dipende dall'organizzazione che si vuole adottare) che interagiscono nel deposito di materiale di una stessa disciplina o argomento specifico.

SPARC ha recentemente definito gli archivi istituzionali come "collezioni digitali che catturano e conservano le produzioni intellettuali di una singola o di più comunità accademiche" Mentre gli archivi disciplinari come ArXiv o CogPrint si focalizzano in particolare su domini di soggetto specifici, i depositi istituzionali immagazzinano e rendono accessibili le produzioni della propria istituzione. Sotto questo punto di vista concettualmente non sono necessariamente limitati agli e-prints di letteratura di ricerca. I server istituzionali potenzialmente potrebbero anche ospitare ed offrire un focus di tipo istituzionale completo per l'aggregazione e conservazione di collezioni diverse da quelle riferibili ai soli paper di ricerca, comprendendo anche dati scientifici, risorse didattiche, collezioni di immagini e molti altri differenti tipologie di contenuto.

I depositi istituzionali possono essere considerati estensione naturale delle responsabilità dell'istituzione accademica in qualità di generatori di ricerca primaria e sono potenzialmente la componente più importante nell'evoluzione della struttura dei nuovi modelli di comunicazione scientifica. Le potenzialità di un deposito istituzionale possono essere numerose:

- giocano un ruolo significativo nel processo evolutivo della ristrutturazione della comunicazione scientifica
- offrono una risposta strategica ai problemi esistenti nel sistema che regola i periodici scientifici
- forniscono un immediato complemento di qualità al modello di pubblicazione scientifica esistente
- stimolano l'innovazione entro una struttura di editoria disaggregata
- sono indicatori tangibili della qualità di un'istituzione, incrementandone la visibilità, il prestigio e il valore a livello pubblico

E' ormai indiscusso che per esserci impatto è necessaria un'ampia disseminazione, in altri termini i lavori dei ricercatori devono essere letti, citati e utilizzati da altri ricercatori, solo così raggiungono l'impatto utile ad uno sviluppo collettivo, quell'impatto che consente di creare nuove ricerche, di effettuare nuove scoperte sulla base di un lavoro altrui letto, assimilato, metabolizzato.

Si stimano circa 2 milioni di articoli annui in 20.000 riviste che ora sono tenuti "prigionieri" entro riviste scientifiche a stampa.

Due sono i canali dell'Open Access:

1. Pubblicazioni entro iniziative di *editoria elettronica sostenibile*
2. Deposito negli Open Archives

L'attuale sistema è in forte conflitto con gli scopi di ricercatori e scienziati i quali, è ovvio, pubblicano i propri lavori di ricerca principalmente per ottenere una massimizzazione dell'impatto entro la comunità internazionale. L'impatto consiste nella lettura dei paper scientifici, nella loro citazione da parte di altri studiosi e nella costruzione di nuove ricerche generate dal lavoro "comunicato".

Due sono le strade che facilitano il ribaltamento dell'attuale modello:

1. il supporto da parte delle istituzioni all'editoria sostenibile per la pubblicazione di periodici open access
2. la creazione di depositi alimentati attraverso il processo di auto-archiviazione per l'archiviazione e disseminazione dei lavori di ricerca

Di editoria sostenibile parlerà Paola Gargiulo che ci illustrerà l'iniziativa SPARC *The Scholarly Publishing and Academic Resources Coalition Europe*.

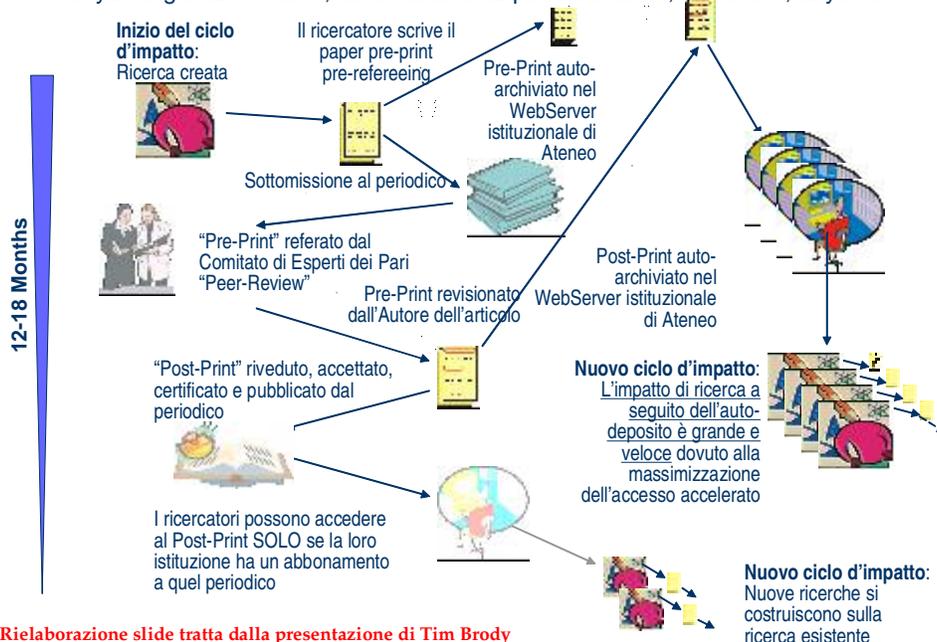
La creazione, gestione, manutenzione, di Open Archives, pur presupponendo un solido e ben definito corpo di "politiche" a corredo dell'archivio o degli archivi, non rientra nella sfera "editoriale". I depositi, siano essi istituzionali, siano essi disciplinari, si comportano in modo differente rispetto alle iniziative dei periodici "open access", pur rientrando e facendo parte di quella zona cosiddetta "open access".

Sostanzialmente un modo per supportare l'open access è quello di continuare a pubblicare nelle riviste tradizionali esistenti munite di peer-reviews, ma di incominciare la pratica dell'auto-archiviazione di quanto sottomesso alle riviste a stampa o on-line. In questo modo si attua un doppio binario; da una parte il deposito immediato entro un archivio di e-print pubblicamente accessibile (istituzionale o disciplinare), dall'altra la sottomissione alla rivista secondo il processo tradizionale, come ben evidenziato da Stevan Harnad e Tim Brody³ nel diagramma di sotto che rappresenta la catena del papero entro una visione nuova di comunicazione scientifica

³ Stevan Harnad, Self-Archive Unto Others as Ye Would Have them Self-Archive Unto You (graphics Tim Brody)
<http://www.ecs.soton.ac.uk/~harnad/Temp/unto-others.html>

Maximized Research Access and Impact Through Self-Archiving

“Skywriting”: All research, accessible to all potential users, anywhere, anytime



Gli ostacoli al diritto di accesso ai contenuti possono essere correlati a due fattori strettamente correlati:

- al prezzo di abbonamento della rivista, sempre più alto e che comunque la rende inaccessibile a molti
- ai diritti di proprietà intellettuale relativi ai contenuti ceduti agli editori.

Relativamente al primo ostacolo le considerazioni che si possono fare sono molteplici. Un prezzo di abbonamento alto rende la rivista scarsamente presente nelle biblioteche; laddove la rivista non è distribuita in modo capillare ne consegue una perdita di efficacia dei contenuti pubblicati in quella rivista, anche se questi sono indicizzati dai maggiori database. Il fatto che una rivista coi suoi lavori sia coperta da autorevoli database specializzati non necessariamente significa che i lavori in essa contenuti vengano letti se il prezzo d'abbonamento risulta elevato. Il fattore d'impatto può rimanere basso, a causa di un prezzo che tende a salire.

Ci dice Giuseppe Vitello⁴, che *“mentre gli editori di varia hanno (o pretendono di avere) margini di profitti che nella migliore delle ipotesi raggiunge il 10% del prezzo di copertina, gli editori scientifici dichiarano costantemente margini a due cifre, che in qualche caso raggiungono il 30-40% del loro fatturato”*, come nel caso di John Wiley dove il margine di profitto arriva al 48.1% o Elsevier con il 38.6%. E' innegabile che il “modello economico” di questi editori, se risulta catastrofico per la ricerca, per il mercato economico invece è un “buon” modello su cui “investire” in quanto fonte di profitto. Per questa ragione risulta assai

⁴ Giuseppe Vitello (2003), Editori «alternativi» e biblioteche digitali nell'economia della comunicazione scientifica. Prossima pubblicazione su Biblioteche Oggi (mese di giugno 2003)

difficile che le attuali leggi sul copyright, pesantemente orientate alla tutela di un mercato economico piuttosto che di un mercato culturale delle idee, tengano conto di questi aspetti e perciò allo stato attuale gli ostacoli normativi diminuiscono la visibilità potenziale dei ricercatori e del loro lavoro, riducendo il loro impatto per il fatto che non sempre le biblioteche sono in grado di pagare gli abbonamenti o i diritti di accesso per riviste elettroniche altrettanto costose.

I ricercatori non traggono "guadagni di pubblicazione" grazie alla vendita dei loro articoli pubblicati nelle riviste scientifiche, ma ottengono "guadagni di impatto" se i loro articoli vengono diffusi in modo adeguato. L'obiettivo dei ricercatori è quello che i loro lavori vengano visti dal maggior numero di persone, indipendentemente da questioni "economiche" intese in termini di guadagno economico personale. L'obiettivo principale va quindi oltre la diffusione, ma si estende alla disseminazione informativa dei loro contenuti intellettuali.

Come ci dice Stevan Harnad nella "teoria sovversiva" da lui elaborata per abbattere i costi delle riviste e rompere questa catena pesantissima, *"i costi elevati nell'era cartacea di Gutenberg, dispendiosa e inefficace, erano inevitabili; ma oggi, nell'era post-Gutenberg on line, il funzionamento alla vecchia maniera, con i suoi costi elevati deve essere mantenuto come opzione complementare invece che come strumento indispensabile."*

Vale a dire: ribaltare la situazione, preferire la versione online e mantenere il cartaceo solo come opzione. Ma Harnad non intende per versione online quella venduta come "prodotto confezionato" dalle case editrici, le quali aggiungono ulteriori sovrapprezzi a quanto già da noi pagato a cifre vertiginose. Quello a cui ci riferiamo (gruppo Open Archive Initiative) è la versione libera dei lavori posti sugli e-print server degli Atenei in un'ottica di "comunicazione aperta" e di interoperabilità tra comunità differenti. Attenzione alle idee di 'valore aggiunto' proposte dagli editori come indispensabili. Secondo la teoria di Harnad il solo servizio indispensabile ancora fornito dagli editori di riviste è la valutazione degli articoli. Gli altri servizi a valore aggiunto quali versione cartacea (se teniamo conto di quella online come opzione preferibile), riproduzioni di pagine in formato PDF on line, sono *optional* di lusso, indubbiamente interessanti, ma non prioritari.

Nell'era digitale non vi è più necessità alcuna di tenere la ricerca valutata prigioniera dei diritti di abbonamento o di considerarla alla stregua degli optional di lusso, i quali vengono pagati attraverso questi diritti. Non vi sono giustificazioni per mantenere questo stato di cose.

Le proposte di sostituzione dell'abbonamento con un diritto di accesso in licenza d'uso, o del diritto di accesso con un pagamento per la selezione di articoli che si legge, pay-per-view e così via... sono tutte modalità che ripropongono lo stesso "modello economico distorto" e su queste nuove vie di accesso all'informazione andrebbero poste serie riflessioni.

I diritti di abbonamento, se eccessivi, sono ostacoli all'accesso e dunque all'impatto ed a una libera ricerca. L'abbonamento è il mezzo per un editore di una rivista di rifarsi delle spese e di garantirsi un profitto equo, ma spesso non si tratta di un profitto equo.

E' tempo ormai che ogni ricercatore si ponga nell'ottica di fare in modo che certi meccanismi distorti vengano riequilibrati a favore della ricerca a tutela del suo stesso lavoro, ma anche degli interessi dell'Università in cui opera. Questo è già possibile con l'auto-archiviazione negli archivi elettronici.

E' importante sottolineare che, oltre alle regole del copyright, che dipendono dal diritto (internazionale, europeo, nazionale) certe riviste hanno regole di embargo, che dipendono semplicemente dalla loro politica editoriale (e non dal diritto).

Vi sono alcuni editori, per esempio che, invocando la *Regola di Ingelfinger*⁵ (sull'embargo), che serve agli editori per proteggere le loro fonti di guadagno, decretano che non valuteranno e ancora meno pubblicheranno, articoli che sono già stati 'resi pubblici' a convegni, o già auto-archiviati nei server delle istituzioni.

Alcune riviste - tra cui Science - che applicavano la *Regola di Ingelfinger* stanno rivedendo le loro posizioni. Nature per esempio l'ha già eliminata, in quanto essa è in contraddizione con gli interessi della ricerca e dei ricercatori e pertanto nuoce all'immagine stessa della rivista.

I ricercatori farebbero bene a ignorare completamente norme che non trovano la minima giustificazione nel diritto, così come hanno fatto da dieci anni a questa parte gli autori dei 200.000 articoli del server ArXiv di fisica.

I contratti di cessione dei diritti economici possono essere stilati in modo personalizzato, e l'autore deve poter intervenire in questa delicata fase.

Laddove nei contratti di cessione non si autorizzino esplicitamente l'auto-archiviazione on line, l'autore dovrebbe modificarne la formulazione in modo che essa venga autorizzata.

Uno schema di clausola che potremmo intitolare "Clausola del Fair Use" può essere così formulata:

*"Delego a [l'editore o la rivista] tutti i diritti di vendere o di cedere il testo (su carta o on line) del mio articolo [titolo dell'articolo]. Conservo solo il diritto di diffonderlo gratuitamente a fini scientifici o didattici, in particolare il diritto di auto-archiviarlo pubblicamente on line sul Web."*⁶

Il 10% degli editori autorizza fin d'ora esplicitamente l'auto-archiviazione delle pubblicazioni valutate e molti editori (Harnad stima attorno al 70%) accetteranno la clausola se inserita nel contratto.

Il progetto europeo RoMEO (*Rights on METadata for Open archiving*)⁷, (1 agosto 2002 - 31 luglio 2003) sorto per investigare la questione dei diritti che ruotano attorno l'auto-archiviazione per i lavori di ricerca depositati negli Open Archive della comunità accademica britannica, ci sta fornendo dei dati ancor più ottimistici.

Su 7.135 periodici pubblicati da 80 editori accademici, ben 3.897 pari al 54.6% supportano formalmente il self-archiving dei pre-print o dei post-print o entrambe le versioni. Solo pre-print 2.552 periodici pari al 35.7%, solo post-print 139 titoli pari all'1.9%, entrambe le versioni 1.209 periodici distribuiti da 20 editori. Rimangono fuori 3.238 periodici pari al 46.57%; va detto però che può anche essere che questi editori non si siano ancora posti la questione dell'auto-archiviazione e comunque non è escluso che, come suggerisce Harnad, qualora l'autore rivendicasse i suoi diritti questi editori non concedano l'autorizzazione all'auto-archiviazione.

Il progetto RoMEO tiene conto delle questioni correlate alla raccolta dei metadati dai data providers entro i service providers entro il quadro OAI. Romeo è coordinato da Charles Oppenheim dell'Università di Loughborough, e si occupa di indagare come i metadati e i dati (lavori originali di ricerca) *'give-away research literature'* sono utilizzati e come questi possono essere protetti. Saranno sviluppati una serie di elementi sui diritti a partire dalla costruzione da esistenti schemi e vocabolari (Open Digital Rights Language) da inserire entro i metadati. Obiettivo di Romeo è creare un sistema

⁵ Harnad, Stevan (2000) Ingelfinger Over-Ruled: The Role of the Web in the Future of Refereed Medical Journal Publishing. *The Lancet Perspectives* 256 (December Supplement): s16.

<http://cogprints.ecs.soton.ac.uk/~harnad/Papers/Harnad/harnad00.lancet.htm>

⁶ formula suggerita da Harnad

⁷ Informazioni sul Progetto Romeo

(Rights METadata for Open archiving)

<http://www.lboro.ac.uk/departments/ls/disresearch/romeo/index.html>

che dimostri come i diritti sui metadati possono essere assegnati, scoperti, raccolti e visualizzati agli utenti attraverso il Protocollo OAI per la raccolta dei metadati.

Gli archivi elettronici dovrebbero essere tutti concepiti entro un quadro di interoperabilità. La costruzione degli archivi si basa solitamente su software libero. Gli Open Archive possono essere resi ricercabili come se facessero parte di un archivio virtuale globale con tutta la letteratura scientifica disponibile, pre e post valutazione. A tal fine ogni nuova installazione, per essere visibile e ricercabile a livello globale, deve essere opportunamente registrata nell'apposito registro internazionale presso il sito OAI Open Archives Initiative⁸.

In tale ottica è importante segnalare alcune iniziative, suddivise in due aree distinte ma strettamente in relazione.

Da una parte si trovano iniziative come OAI la quale è nata con lo scopo di sviluppare e promuovere standard per l'interoperabilità tra archivi al fine di essere di supporto ad un'efficiente disseminazione di contenuti; dall'altra si collocano iniziative come la Budapest Open Access Initiative BOAI o il movimento Free Online Scholarship FOS.

La *Open Archive Initiative OAI*⁹ si occupa di approntare standard e protocolli per l'interoperabilità tra archivi, per questo si parla di compatibilità OAI. OAI fonda le sue radici nel continuo sforzo teso al miglioramento dell'accesso ai documenti entro gli archivi e-print, mettendo a disposizione strumenti, software e documentazione (protocolli, standard) e promuovendo l'utilizzo di metadati standard. OAI è supportata nelle sue iniziative dalla Digital Library Federation e da Coalition for Networked Information e riceve finanziamenti da National Science Foundation.

Nella terminologia *Open Archive Initiative (OAI)*, gli e-print server sono detti "data provider" o "fornitori di dati", laddove per dati si intendono sia i "testi pieni" (full-text) sia i "dati sui dati" ovvero i metadati. I metadati sono le informazioni che descrivono i dati, per esempio i dati bibliografici che descrivono un articolo depositato sul server.

In OAI esistono perciò due categorie di server:

- i data provider, DP, i quali contengono i documenti depositati e che supportano il protocollo OAI esponendo i metadati relativi ai lavori contenuti nel server. Un data provider gestisce uno o più server, per il deposito (Web servers repositories)
- i service provider, SP, usano i metadati dei data provider come base per la costruzione di servizi a valore aggiunto.

L'iniziativa di Budapest si distingue da OAI giacché il suo scopo è di fare pressione sugli autori al fine di ottenere un consenso generalizzato ed è finanziata dalla Fondazione Soros di OSI Open Society Institute¹⁰. OSI elargirà un milione di dollari per anno, per tre anni consecutivi, per lo sviluppo di un progetto il quale prevede piani e modelli per l'auto-archiviazione "sostenibile" in oltre 140 Paesi. Tra gli ideatori di BOAI ci sono nomi come Fred Friend: (Director Scholarly Communication, University College London), o Rick Johnson (Director, Scholarly Publishing and Academic Resources Coalition (SPARC) e anche rappresentanti del mondo dell'editoria, Jan Velterop di BioMed Central per esempio. Per BOAI ciò che conta è il consenso dell'autore, non soltanto le necessità o desideri

⁸ <http://www.openarchives.org/data/registerasprovider.html>

⁹ Antonella De Robbio *Open Archives Initiative (OAI) in Europa: Workshop al CERN di Ginevra*.

In "Biblioteche Oggi". Maggio 2001, pp. 66-69

¹⁰ OSI è nata nel 1993 dal sostegno finanziario di Soros con lo scopo di offrire supporto allo sviluppo di programmi (sociali, educativi, per i diritti delle donne, ...) nei Paesi in via di sviluppo: Europa dell'Est, ex Unione Sovietica, Haiti, Mongolia, Africa, America Latina, ...

dell'utenza. BOAI non si occupa di applicazioni software o di aspetti tecnologici, ma è focalizzata principalmente sui processi di peer-reviews. Il suo scopo primario è quello di promuovere l'auto-archiviazione.

Il concetto che sottende la filosofia BOAI è quello di un rigoroso rispetto del diritto d'autore, che libera la ricerca. E' un concetto molto vicino a quello del software libero distribuito con licenza GNU, di cui parlerò in seguito.

Va sottolineato però che non è sufficiente creare degli archivi elettronici, ma è necessario che tutti i ricercatori comincino ad auto-archiviare i loro articoli.

Le biblioteche e i bibliotecari possono essere un punto di riferimento, un supporto organizzativo all'archiviazione stessa dei documenti in stretta sinergia con gli informatici.

Fondamentale è l'organizzazione di service provider dedicati alla raccolta di metadati dai vari archivi, o sistemi di metadata harvesting¹¹, entro una cornice di interoperabilità estesa al di là delle differenze tra server. Tale funzione inquadrabile come servizio a valore aggiunto, è una funzione di meta-livello, che sta sopra gli Open Archive intesi come Data Provider distribuiti sul territorio. Va detto che, nell'implementare un server OAI, sarebbe opportuno attenersi al protocollo Open Archives Initiative Protocol for Metadata Harvesting Protocol¹² indispensabile per rendersi visibili ai fini della raccolta, esponendo i propri metadati al colloquio entro la comunità internazionale.

A livello internazionale lo scenario si svolge attorno a due canali, apparentemente distinti, ma di fatto interagenti l'uno con l'altro. Gli archivi istituzionali da una parte, nuova frontiera, numerosi, ma piccoli e poco "popolati" e, secondo alcune scuole di pensiero tra cui quella di Stevan Harnad, a livello politico con una forza strategicamente innovativa. Gli archivi disciplinari dall'altra, più ampi, che rappresentano momenti di forte aggregazione per alcune comunità scientifiche, meglio collaudati e di cui si registrano alcune iniziative ben avviate che riflettono realtà ormai consolidate. Esistono inoltre altre soluzioni: gli "umbrella server" (specie di MetaOA), individual data provider (modello Kepler), ...

I fisici già da dieci anni archiviano i loro lavori nei server di tutto il mondo, furono i primi a partire con il server xxx.lanl.gov ora arXiv.org. I ricercatori in scienze cognitive in tre anni hanno depositato oltre mille articoli¹³.

Il progetto degli economisti WoPEc (Working Papers in Economics), entro la rete RePEC, sorto nel 1999 per opera di Thomas Krichel¹⁴ è costituito da una rete di oltre 200 archivi¹⁵ che mettono a disposizione più di oltre 60.000 articoli e rapporti tecnici di ambito economico disponibili online. Ogni istituzione che partecipa al progetto mette a disposizione i propri dati ma, contemporaneamente, diventa un *mirror* che garantisce l'accesso anche ai dati di tutti gli altri archivi.

¹¹ Per esempio: ALCME: OAIHarvester Project

<http://alcme.oclc.org/OAIHarvester.html>

¹² Protocol Version 1.1 of 2001-07-02 messo a punto da Herbert Van de Sompel, Carl Lagoze

Document Version 2001-06-20

<http://www.openarchives.org/OAI/openarchivesprotocol.htm>

¹³ CogPrints Archive The Cognitive Sciences E-Print Archive

<<http://cogprints.soton.ac.uk/>>

Computer Science and Engineering, Psychology, Neuroscience, Behavioral Biology, Linguistics and Philosophy

¹⁴ economista ora docente di biblioteconomia a New York

¹⁵ Per la lista degli archivi vedere:

<http://ideas.uqam.ca/ideas/archives.html>

Anche i matematici hanno sviluppato progetti internazionali per la creazione di *repositories* dei loro lavori¹⁶. Tali progetti sono per esempio *Front for the Mathematics*, con oltre 10.000 lavori matematici¹⁷, *ArXiv*¹⁸ entro il server dei fisici di Los Alamos. Il progetto europeo *MPRESS/MathNet.preprints server*¹⁹, coordinato dai matematici tedeschi Judith Plümer, Roland Schwänzl, non è un server per il deposito, ma un'architettura detta di tipo "*umbrella server*", la quale consente di ricercare nei metadati di alcuni server di preprint, attraverso una meta-interfaccia.

Per la medicina esistono esperienze nuove, iniziate a partire dalla fine del 2000, in connessione con prestigiose banche dati, come nel caso dell'integrazione tra la banca dati *PubMed (Medline)* e il repository *PubMedCentral*²⁰, tipo di archivio centralizzato che funziona attraverso accordi con editori e che prevede un formato XML per il trattamento non solo dei metadati, ma anche dei dati (testi pieni). *PubMedCentral* in sostanza è un server centralizzato, un archivio dove però non avviene *self-archiving*, ma i lavori vengono depositati dagli stessi editori.

*NetPrint*²¹, per il settore della medicina clinica, è molto più vicino ad un vero open archive, in quanto si basa sull'auto-archiviazione da parte degli autori.

Esistono anche Open Archive per la chimica, la biologia e per altre discipline, anche se per quelle umanistiche c'è ancora molta strada da percorrere.

*Cogprints*²² di Stevan Harnad è l'OA per le scienze cognitive, psicologia, neuroscienze e linguistica dove i ricercatori in tre anni hanno depositato oltre mille articoli a testo pieno, entro un archivio che comprende oltre 6500 metadati, recuperati da altri archivi di bibliografie in BibTex. *Cogprint* è un esempio di e-server misto dove convivono appunto soli metadati provenienti da bibliografie e metadati dotati del corrispettivo documento a testo pieno provenienti da processi di self-archiving da parte degli autori.

Esistono inoltre esperienze nel campo dell'informatica, è il caso di ricordare il progetto *NCSTRL* (pronuncia ancestral) *Networked Computer Science Technical Reference Library*²³, a cui fa riferimento anche *ERCIM* per la matematica, il quale sta per traslocare e transitare dal vecchio modello che implementava il protocollo Dienst verso il modello che utilizza tecnologie OpenArchive Iniziative di tipo EPrints entro una cornice di progetto che prenderà il nome di *Scholnet*. La trasformazione di *NCSTRL* coinvolgerà probabilmente anche *ERCIM*, the European Research Consortium for Informatics and Mathematics che ne è parte integrante. *Scholnet* prevede nuovo software per la gestione di materiale multimediale e didattico per attività di e-learning, e per la creazione ed estrazione automatica di metadati.

Lo status quo è una situazione frammentaria con centinaia di e-server istituzionali, disciplinari, accentrati, distribuiti, piccoli, grandi, depositi di collezioni, Papers sparsi sui siti Web, Umbrella servers e MetaMotori OA (service provider).

¹⁶ <http://www.ams.org/global-preprints/umbrella-server.html>

¹⁷ Front End for the Mathematics ArXiv
<<http://front.math.ucdavis.edu/>>

¹⁸ <http://arXiv.org/>

¹⁹ MPRESS - MathNet.preprints
<<http://MathNet.preprints.org/>>

²⁰ PubMedCentral <<http://www.pubmedcentral.nih.gov/>>

²¹ Cincial Medicine NetPrint <<http://clinmed.netprints.org/>>

²² CogPrints Archive The Cognitive Sciences E-Print Archive (Computer Science and Engineering, Psychology, Neuroscience, Behavioral Biology, Linguistics and Philosophy), <<http://cogprints.soton.ac.uk/>>.

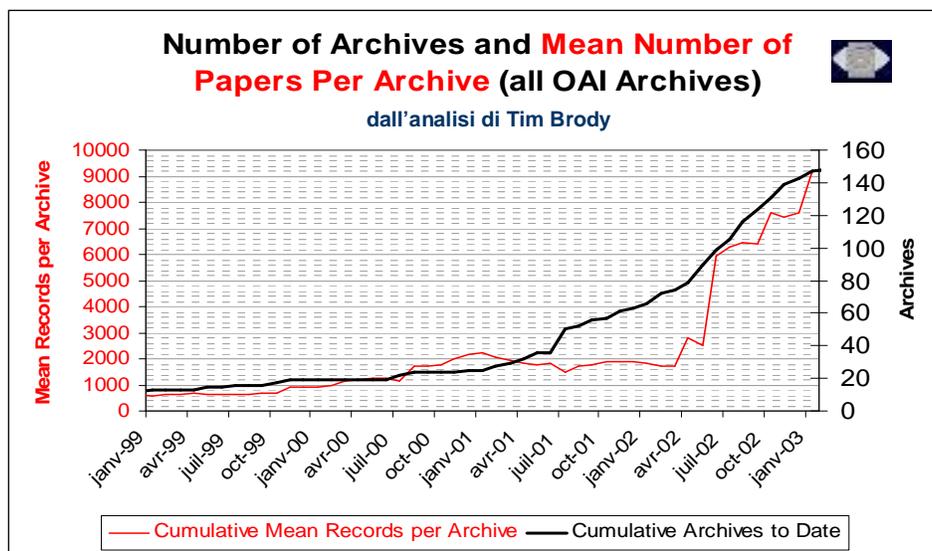
²³ <<http://www.ncstrl.org/>>.

L'esperienza di CiteSeer/ResearchIndex, costruito da NEC Research Institute nasce per sanare questa frammentarietà. Il servizio pesca regolarmente home page degli studiosi nel campo dell'informatica, catturando e collezionando paper. Consente ai singoli autori di depositare i lavori ed è connesso con arXiv attraverso una sezione relativa all'informatica, chiamata CoRR Computing Research Repository. Le esperienze di "umbrella server" del tipo MPRESS sopra citato offrono un utile servizio di meta-ricerca che consente di interrogare più server attraverso metadati Dublin Core, ma tale soluzione non offre un deposito unico.

Un vero modello di sistema di archivi aperti di tipo distribuito è RePEC *Research Papers in Economics*²⁴, la più grande collezione decentralizzata del mondo di documenti per l'economia ad accesso libero, una rete che collega diversi archivi.

Prima di passare ad analizzare gli ostacoli intesi come rischi di progetto nell'apertura di depositi disciplinari o istituzionali e prima di illustrare alcuni suggerimenti per possibili soluzioni, vediamo alcuni dati statistici sugli Open Archive, elaborati da Tim Brody²⁵. Ad oggi esistono oltre 140 archivi, tra disciplinari e istituzionali, di cui 3 di grandi dimensioni (descritti sopra) e ricchi di papers e 68 basati sul modello EPrints.

Crescita del numero di archivi OAI
ad oggi esistono oltre 140 archivi, ma la media del numero di
papers per Archivio (9000) necessità di una crescita veloce



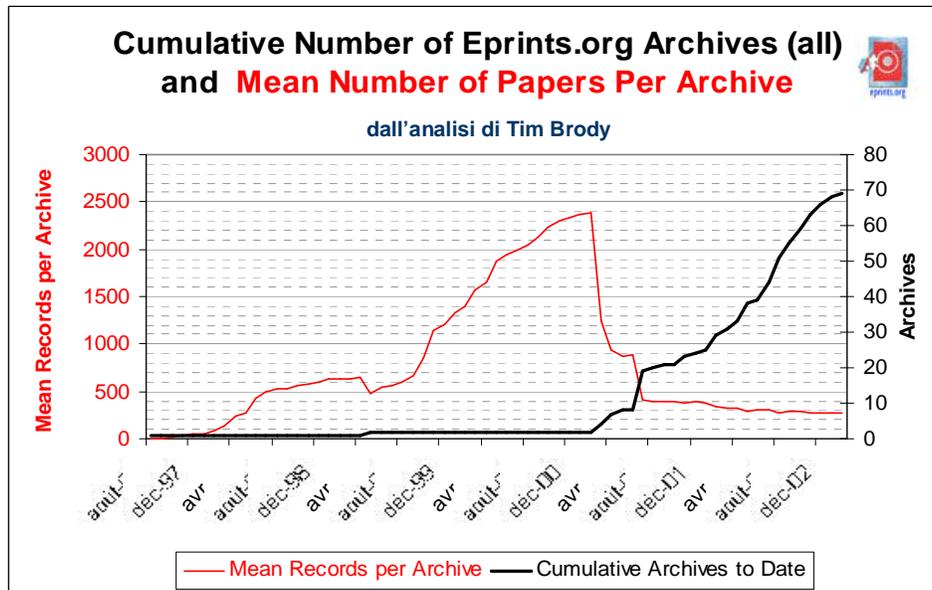
La media del numero totale di depositi per archivio è circa 9000 (tenendo conto dei tre grandi e considerando la media bassa degli EPrints) e va necessariamente innalzata velocemente.

²⁴ <<http://repec.org/>>

²⁵ Tim Brody, Institutional OAI Archive Growth (Data collected and analysed by Tim Brody, doctoral candidate, Electronics and Computer Science, Southampton University)
www.ecs.soton.ac.uk/~harnad/Temp/tim.ppt

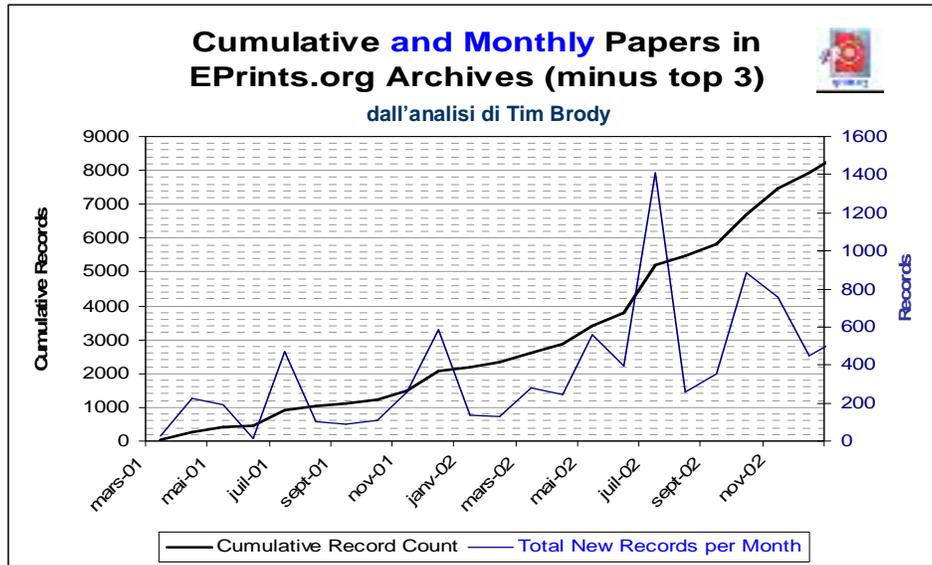
L'incremento del numero di lavori free access entro gli archivi OAI è pari a oltre 1.400.000 record totali (compresi i 234.181 di ArXive e i 200.000 di RePEc), anche se non tutti i record si riferiscono a lavori depositati a testo pieno, ma ai soli metadati.

Il modelloEprints.org sottoinsieme di OAI (sono circa la metà di tutti gli archivi correnti OAI Archives, 70/140)



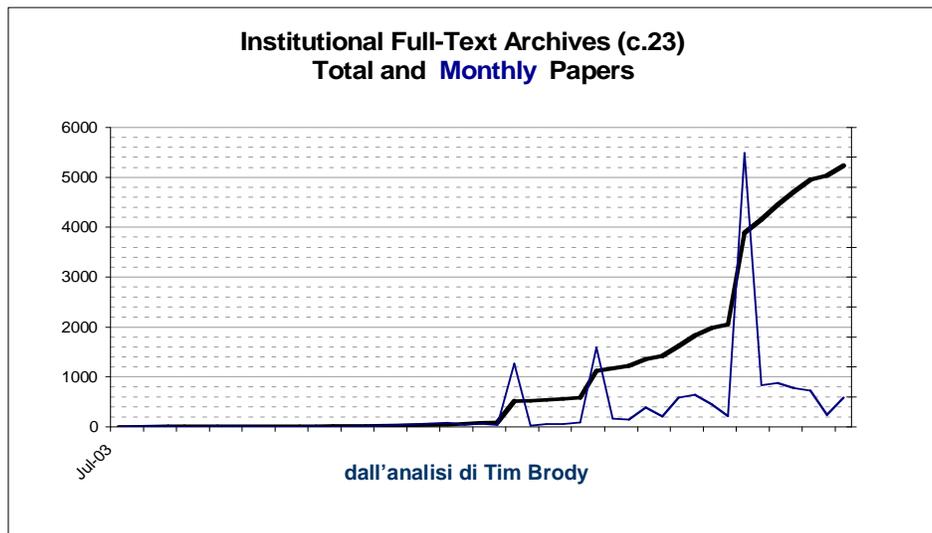
Il totale dei record di tutti i 68 Archivi Eprints è di oltre 8000, con una media di 120 papers depositati in ciascuno (va considerato che il modello EPrints è appena nato).

**Crescita dei papers negli archivi Eprints
(esclusi i tre grandi)
oltre 8000 papers ad oggi (maggio 2003)**



Gli Archivi istituzionali assestati e operativi nel mondo sono 23 con un totale di 5000 papers full-text (processo lento)

**Crescita nel numero di papers a testo pieno depositati negli archivi istituzionali
(oltre 5000 papers in 23 archivi)**



Aprire un Open Archive non è un problema tecnico, ma è una questione di “politiche” a vari livelli. Il deposito deve essere considerato il registro ufficiale delle produzioni di ricerca di quell’ateneo. Deve essere riconosciuto come strumento di ausilio alla valutazione delle produzioni di ricerca degli studiosi di quell’ente da parte degli organismi preposti alla valutazione. Dobbiamo chiederci “Che cosa serve per riempire gli Open Archives?”

L’obiettivo primario delle Università (o degli enti finanziatori della ricerca) deve essere quello di adottare una politica sistematica di auto-archiviazione di tutte le produzioni di ricerca “*refereed*”.

Servono *policies* a corredo dell’Open Archive *per la submission, di copyright, e sui metadati*.

Altro passo fondamentale è quello di adottare Curricula Vitae CV online standardizzati, sul modello britannico²⁶, questo potrà essere uno strumento in più per la valutazione online dell’impatto della ricerca da parte di chi finanzia la ricerca, attraverso opportune connessioni tra i CVs e i dati entro gli archivi.

Senza il consenso non si apre nessun Open Archive, né disciplinare, né istituzionale. Ma occorre comprendere bene quale dovrà essere il modello organizzativo più congeniale, se appunto un deposito istituzionale oppure più depositi disciplinari entro lo stesso ente, oppure condivisi da più enti. Non una questione puramente tecnica perché gli strumenti software per costruire un archivio ci sono e peraltro sono strumenti facili da utilizzare, da gestire, sono OpenSource. Serve il consenso degli organi accademici. Servono le politiche, per la sottomissione dei dati e dei metadati, le politiche di copyright, serve stabilire chi sarà l’entità Open Archive di quell’ente, lo si deve definire in tutte le sue linee. Occorre dargli un’identità precisa, scegliere un nome, stabilire chi può accedere, chi può depositare, che cosa va depositato e cosa no, quali sono i formati accettati. E’ necessario, a scanso di pesanti equivoci, stabilire se deve essere un deposito che contenga “tutto” quello che viene prodotto dall’ente, pre-print, report tecnici, capitoli di libri, interventi a seminari, presentazioni ... o solo ciò che viene sottomesso a periodici o presentato a convegni.

Ma attenzione, un deposito non deve essere una rivista, sono due strumenti diversi e questo va detto agli organi che dovranno deliberare l’apertura di un open archive entro la struttura. Ai bibliotecari va detto che un archivio non è un catalogo, è uno strumento fondamentalmente diverso e quindi sia l’interfaccia sia i metadati devono essere “a portata di mano” dell’utente finale, così come lo schema disciplinare per il browsing utile alla “collocazione” entro uno schema disciplinare o anche puramente gerarchico che dovrà essere il più semplice possibile. Questo perché poi avremo bisogno soprattutto del consenso degli autori, altrimenti ogni tentativo di far decollare un open archive fallirà. In Italia attualmente abbiamo qualche esperienza, ma i server sono vuoti, appunto perché manca il consenso.

La chiave del successo sta nell’aggregazione e nella cooperazione.

Fondamentalmente si tratta di un problema economico e le soluzioni stanno nelle potenzialità del lavoro distribuito e nella creazione di alleanze, cruciali nella creazione di reti per gli Open Archive.

Per tale ragione è vincolante per una buona riuscita di qualsiasi che alla base vi sia una politica sul copyright basata su:

²⁶ Template and demo for a standardized university online-CV with harvestable performance indicators and links to the institutional Eprint Archive

http://paracite.eprints.org/cgi-bin/rae_front.cgi

- linee guida chiare di supporto agli open archives italiani con indicazioni su come stendere le politiche di copyright a corredo degli archivi, documenti ad opera di un Working Group italiano (tutto da inventare)
- azioni di promozione alla pratica dell'auto-archiviazione
- diffusione dei risultati che emergeranno dal lavoro del gruppo europeo RoMEO Project (Rights METadata for Open archiving)

In merito agli archivi istituzionali i responsabili dovranno vegliare affinché siano osservate alcune precauzioni, offrendo accordi espliciti con gli autori che auto-depositano, per esempio sottoponendo le regole dell'archivio come parte integrante del processo di submission del paper. Queste dovrebbero includere garanzie sul fatto che l'autore non sta violando accordi presi con terze parti, come editori, di fatto che non sta violando il copyright depositando il lavoro. Tali garanzie dovrebbero anche assicurare che l'autore esplicitamente accetta i termini e le condizioni del contratto tra autore e archivio ovvero accetta che il contenuto sia reso liberamente disponibile pubblicamente.

Fino ad oggi l'attenzione sul copyright delle università si è focalizzata sulle questioni IPR (Intellectual Property Rights) sui marchi, brevetti, scoperte industriali o sui materiali di ambito pedagogico, quali i test e molte delle politiche sul copyright per ora escludono gli articoli e i volumi. Alcune università nel mondo si stanno però ponendo il problema della proprietà intellettuale sui lavori di ricerca dei propri ricercatori, e si stanno chiedendo a chi spetti il copyright, se all'ente o se all'autore.

Per convincere gli autori è necessario il consenso al fine di ottenere il deposito di "contents, contents, contents" (grido di Harnad).

E' perciò necessario analizzare i comportamenti e le abitudini delle singole comunità di studiosi: a livello di gruppo, nazionale, internazionale, dentro l'istituzione, nella comunità disciplinare... Come ho esposto sopra, i comportamenti tra le varie comunità differiscono enormemente. Avere un quadro preciso dei comportamenti delle varie comunità di utenti è strategico in quanto le abitudini e le trasformazioni sociali sono più importanti di qualsiasi configurazione tecnologica.

Potrebbe essere importante avviare un'analisi dei comportamenti delle varie comunità, a livello territoriale, ciò al fine di individuare eventuali aperture di archivi disciplinari su base italiana.

La comunità scientifica deve essere coinvolta, per questa ragione, nei dibattiti internazionali, qualcuno richiama l'attenzione sugli approcci basati su disciplina i quali non sono affatto in contraddizione con gli archivi di tipo istituzionale.

Per usare tecnicismi mutuati dai settori LIS, potremmo parlare di indagine preventiva sugli "user requirements" per la costruzione di Archivi confezionati sulle reali esigenze delle comunità (locali? Centralizzati?). Non basta dire che online è meglio (il grido di Krichel), occorre offrire qualcosa in più. Gli archivi istituzionali devono essere integrati dentro la Biblioteca Digitale in modo da consentire una visibilità totale e trasversale da qualsiasi punto del sistema informativo: soluzione portale, OPAC, banche dati, periodici elettronici, e in modo da garantire una connessione tra referenze bibliografiche.

Le parole chiave per il consenso e che aggiungono valore all'atto dell'auto-archiviazione sono:

- ricerca a testo pieno sui documenti depositati: applicazione delle funzionalità dei motori di ricerca per le ricerche nei full-text: vedi esperienze CERN e Caltech
- analisi delle transazioni log: statistiche sugli accessi al server, ma anche sul singolo autore e su ogni suo singolo articolo (vedi esempio ArXive, RePEc per abstract visualizzati, articoli scaricati, per giorno, mese, anno ...)

- Analisi scientometriche e webmetriche: utilizzo e applicazione di nuovi strumenti per il calcolo e modalità nuove per la misurazione dell'impatto di ricerca, per il recupero e linking delle citazioni nel web (vedi Paracite, OpCite, CiteBase, Citeseer/ResearchIndex)

Il Reference Linking è al centro di importanti ricerche in alcuni settori di ricerca in campo bibliometrico, in particolare nell'analisi citazionale.

Reference Linking and Citation Analysis combinati assieme compongono un punto di partenza fondamentale entro il quadro Open Archive Initiative per nuovi modelli di comunicazione scientifica.

Esiste un'ampia gamma di lavori, di ragguardevoli proporzioni rispetto alla totalità delle produzioni intellettuali scientifiche, che non ricevono citazioni nei successivi cinque anni dalla loro pubblicazione. Questa zona chiamata *area uncitedness* è ugualmente degna di nota per la sua breve ma tormentosa storia la quale è soggetta ad ambiguità e fraintendimenti, errori di calcolo e politicizzazioni. In un articolo del 1997 Schwartz²⁷ rivaluta questa grande massa di uncitedness in termini di analisi di un fenomeno generale insito nel circuito della comunicazione scientifica. Le percentuali dei lavori che ricadono in zona uncitedness variano a seconda degli ambiti disciplinari, raggiungendo per l'ambito LIS il 72% del totale dei lavori prodotti.

L'applicazione di strumenti per il reference linking e il parsing citazionale entro gli Open Archive aiuta a ridurre la zona uncitedness facendo emergere lavori anche importanti che restano nascosti dopo la loro pubblicazione formale in riviste a stampa.

I servizi Open Archive di CiteBase sono una ricca fonte di informazione e un punto di partenza verso il movimento dell'auto-archiviazione. CiteBase mette a disposizione il software per la costruzione di archivi (EPrints), il software per il parsing citazionale OpCit ai fini della connessione tra citazioni (reference linking). CiteBase Search, disegnato e scritto da Tim Brody, fa parte del progetto Open Citation Project, fornisce agli utenti la possibilità di ricercare entro archivi multipli differenti con funzionalità di ranking nella presentazione dei risultati, incluso anche l'impatto della citazione. I dati sono raccolti usando il protocollo Open Archives Metadata e un harvester che si occupa di scaricare i dati da un archivio all'altro, in questo caso da ArXiv e da Cogprints, l'archivio aperto per le scienze cognitive.

ParaCite è un servizio sperimentale progettato sempre dal gruppo dall'università di Southampton, per la localizzazione degli articoli dai riferimenti bibliografici grezzi. Quando una citazione bibliografica è passata al servizio, viene suddivisa in parti componenti (per esempio autore, titolo, anno) e trasferita alla risorsa di ricerca. Viene presentato un insieme di risorse basate sull'area di soggetto e sui dati forniti facendo in modo che il sistema scelga un link verso risorse con alta probabilità accesso all'articolo a testo integrale a costo zero.

ParaCite ha sviluppato un'applicazione per il software EPrints, il pacchetto ParaTools che include la conversione sia di liste di riferimenti bibliografici, sia di metadati entro OpenURL validi, la collezione di metadati dalle bibliografie per realizzare ricerche interne, l'estrazione automatica della bibliografia dai documenti, la realizzazione di ricerche usando ParaCite.

Si tratta di una procedura ancora in fase di test e in via di sviluppo (siamo ancora alla versione 1.0 di ParaTools, disponibile solo dal 29 gennaio 2003) ma per il momento pare assai utile dotare i metadati dei riferimenti bibliografici connessi a servizi come Google, ResearchIndex e CiteBase.

²⁷ Charles A. Schwartz (1997). "The rise and fall of uncitedness," *College & Research Libraries* 58: (January) http://www.ala.org/Content/NavigationMenu/ACRL/Publications/Journals_Monographic_Series/College_and_Research_Libraries/Back_Issues_1997/January97/CandRL_January_1997_abstracts.htm

ResearchIndex, già citato sopra, non è un semplice motore di ricerca, ma pur avendo al suo interno un motore per le ricerche e molti altri strumenti, è un sistema che mette a disposizione tools per la costruzione di biblioteche digitali. Il sistema ACI Autonomous Citation Index al suo interno consente l'indicizzazione automatica con estrazione dal testo pieno di parole significative che vengono contestualizzate e rendono operativo il linking tra le risorse organizzate entro le sue directory.

L'importanza di ResearchIndex entro un Open Archive si esplica soprattutto nel fatto di essere uno strumento per l'analisi citazionale dei documenti, alternativo al costoso SCI dell'ISI, Science Citation Index. Pur essendo ancora agli esordi il suo database sta crescendo vertiginosamente in quanto alimentato dalla comunità degli stessi ricercatori. Per tale ragione una connessione tramite ParaTool da un Open Archive verso ResearchIndex può risultare assai allettante.

Un altro potenziale problema è correlato al destino degli eprints ovvero di una loro conservazione (con tutti i loro contenuti) a lungo termine. Gli accademici costruiscono le loro ricerche sui lavori di altri e regolarmente citano o fanno riferimenti alla passata letteratura. La citazione di riferimenti valida molte delle rivendicazioni che gli scienziati faranno nei loro lavori e li incorporano nel consenso pre-esistente. Uno dei ruoli primari dei periodici a stampa è stato quello di essersi evoluti fino a divenire un archivio di pubblico dominio ben stabilizzato e con funzioni di registro pubblico delle scoperte scientifiche. Ciò era dovuto non tanto ad una responsabilità dell'editore, ma perché le biblioteche di ricerca, collettivamente e attraverso i loro cataloghi, hanno reso un ottimo servizio che agiva appunto come deposito distribuito delle forme cartacee conservando a futura memoria le conoscenze racchiuse nei periodici per gli scienziati presenti e futuri.

Il passaggio verso i contenuti acquistati attraverso licenze minaccia il ruolo delle biblioteche come conservatori delle conoscenze scientifiche. Per esempio la minaccia potrebbe includere la possibilità che l'istituzione che mette in piedi un archivio possa non essere sempre attenta alle proprie responsabilità di assicurare la conservazione a lungo termine dei contenuti.

Per tale ragione sarebbe importante cominciare a riflettere sulle tecniche e sulle strategie di migrazione dei dati o di emulazione, per una conservazione del digitale (*digital preservation*) dei contenuti e dei metadati entro gli Open Archives italiani.

In aggiunta alle questioni pratiche già considerate che possono ostacolare la crescita e lo sviluppo degli archivi, ci sono molti altri aspetti che provengono dal contesto culturale e scientifico di ciascuna comunità. In particolare vanno considerate le funzioni multiple che i periodici peer-reviewed attualmente giocano nel sistema della comunicazione scientifica. Mentre molti fautori dei sistemi eprints, tra cui Harnad, sostengono che gli autori che scrivono sui periodici con comitato di refereeing lo fanno ai fini dell'impatto delle loro ricerche, ci sono innumerevoli sfaccettature nelle motivazioni degli autori e ciò si evince dai multipli ruoli che i periodici hanno avuto nel tempo, evolvendosi e adattandosi ai contesti. Sostanzialmente andrebbero indagate almeno quattro aree motivazionali che intervengono nella scrittura di un lavoro scientifico:

1. il ruolo dei periodici nell'attuale sistema di comunicazione scientifica
2. le differenti motivazioni che portano a scrivere e a pubblicare un lavoro
3. le differenze culturali tra discipline differenti
4. la diversa natura dell'istituzione di ricerca

Odlyzko²⁸ afferma che mentre gli scienziati possono anche essere avventurosi a livello intellettuale, tendono ad essere conservativi nelle abitudini di lavoro, soprattutto rispetto alle nuove tecnologie. La tecnologia non basta a stravolgere i processi della comunicazione scientifica, vanno considerate le motivazioni, che possono essere innumerevoli e talvolta imprevedibili, spesso legate al contesto locale e più in là al contesto di comunità. I rapporti di forza scientifici, ci ricorda Bourdieu²⁹, si esercitano attraverso rapporti di conoscenza e di comunicazione e questo particolare rapporto di potere scientifico di tipo simbolico agisce soltanto su quelle categorie di percezione in grado di conoscerlo e riconoscerlo.

Riferimenti

Liste di discussione internazionali

- Nature: <http://www.nature.com/nature/debates/e-access/index.html>
- Science: <http://www.sciencemag.org/cgi/eletters/291/5512/2318b>
- American Scientist: <http://amsci-forum.amsci.org/archives/september98-forum.html>
- FOS Newsletter and Forum: <http://www.earlham.edu/~peters/fos/index.htm>
- Colloque virtuel text-e: <http://www.text-e.org/>
- SPARC Institutional Repositories Discussion List Archive - SPARC-IR@arl.org:
<https://mx2.arl.org/Lists/SPARC-IR/>
- CNI-COPYRIGHT Forum: <http://www.cni.org/forums/cni-copyright/cni-copyright.html>

Siti di riferimento

- Bozza delle politiche di autoarchiviazione (preprint e postprint) per le Università
<http://www.ecs.soton.ac.uk/~lac/archpol.html>
- Proposta UK per una politica nazionale di auto-deposito nelle Università britanniche di tutte le produzioni di ricerca “refereed” ai fini della valutazione
<http://www.ecs.soton.ac.uk/~harnad/Temp/Ariadne-RAE.doc>
- Template e demo per CV universitari online standardizzati con indicatori di performance esposti agli harvester e links agli Eprint Archive istituzionali
http://paracite.eprints.org/cgi-bin/rae_front.cgi
- RoMEO Project (Loughborough) Rights METadata for Open archiving)
<http://www.lboro.ac.uk/departments/ls/disresearch/romeo/>
 - Strumenti per l'Analisi Citazionale (Scientometrica, Webmetrica)

²⁸ Odlyzko, A. (1995) "Tragic loss or good riddance? The impending demise of traditional scholarly journals." *International Journal of Human-Computer Studies*, 42, p. 86.

http://www.dtc.umn.edu/~odlyzko/doc/tragic_loss_long.pdf

²⁹ Pierre Bourdieu. La parola e il potere. L'economia degli scambi linguistici. Guida, Napoli, 1982.

<http://citebase.eprints.org/help/>
<http://opcit.eprints.org/evaluation/Citebase-evaluation/evaluation-report.html>
<http://paracite.eprints.org/>
<http://citeseer.nj.nec.com/cs>

Bibliografia

- Lawrence, S. (2001) *Free online availability substantially increases a paper's impact*. Nature 411 (6837): 521.
<http://www.neci.nec.com/~lawrence/papers/online-nature01/>
- Guédon, Jean-Claude (2002) *In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing*
<http://www.arl.org/arl/proceedings/138/guedon.html>
- Harnad, Stevan (1990) Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry. Psychological Science 1: 342 - 343 (reprinted in Current Contents 45: 9-13, November 11 1991). <http://cogprints.soton.ac.uk/documents/disk0/00/00/15/81/>
- Harnad, Stevan (2001) For Whom the Gate Tolls? How and Why to Free the Refereed Research Literature Online Through Author/Institution Self-Archiving, Now.
<http://cogprints.soton.ac.uk/documents/disk0/00/00/16/39/>
- Suber, Peter (2003). "Removing the barriers to research: an introduction to open access for librarians." *College & Research Libraries News*, 64, 92-94, 113.
<http://www.earlham.edu/~peters/writing/acrl.htm>
- Odlyzko, Andrew M. (1995) "Tragic loss or good riddance? The impending demise of traditional scholarly journals." *International Journal of Human-Computer Studies*, 42, 71-122.
<http://www.dtc.umn.edu/~odlyzko/doc/tragic.loss.long.pdf>