

# Secondo workshop OAI in Europa

## Gaining independence with e-prints archives and OAI

Relazione di Antonella De Robbio  
[documento del 27-10-2002]

"Librarians can (and ought to) help create  
a navigable, worldwide ocean of knowledge, open to all;  
and, like Odysseus,  
they will know how to help negotiate the tricky ebbs and eddies,  
the vortices and the undertows of chaotic knowledge flows  
that necessarily accompany the development  
of a distributed intelligence civilization -  
a civilization open to all that are good enough (excellence),  
and not only to those who can afford it (elites)."

[Jean-Claude Guedon, 'In Oldenburg's long shadow:  
librarians, research scientists, publishers and the control of scientific literature'<sup>1</sup>]

A distanza di un anno e mezzo dal primo incontro europeo sugli Open Archives<sup>2</sup>, si è tenuto, sempre presso il CERN a Ginevra, dal 17 al 19 ottobre 2002, il "2<sup>nd</sup> Workshop on the Open Archives Initiative (OAI): Gaining independence with e-prints archives and OAI", organizzato da LIBER *Ligue des Bibliothèques Européennes de Recherche*<sup>3</sup>, da SPARC *Europe Scholarly Publishing and Academic Resources Coalition*<sup>4</sup> e dalla *biblioteca del CERN*<sup>5</sup>, con la

---

<sup>1</sup> Guedon, Jean-Claude. 'In Oldenburg's long shadow: librarians, research scientists, publishers and the control of scientific literature'. *Creating the Digital Future, Association of Research Libraries Proceedings of the 138th Annual Meeting*, 2001.

<http://www.arl.org/arl/proceedings/138/guedon.html>

<sup>2</sup> Il primo *Workshop on The Open Archives initiative (OAI) and Peer Review journals in Europe*, promosso da Access Division of LIBER (*Ligue des Bibliothèques Européennes de Recherche*) si tenne sempre al CERN di Ginevra dal 22 al 24 marzo 2001

<http://documents.cern.ch/AGE/current/fullagenda.php?ida=a01193>

Per un resoconto in italiano vedasi De Robbio, Antonella. *Workshop on the open archives initiative (OAI) and peer review journals in Europe*. «AIB Notizie», 13 (2001), n. 5, p. 14-15.

<http://www.aib.it/aib/editoria/n13/01-05derobbio.htm>

<sup>3</sup> LIBER è un'associazione non governativa (ONG) per le biblioteche di ricerca in Europa, nata sotto gli auspici del Consiglio d'Europa la quale rappresenta e promuove gli interessi delle biblioteche di ricerca europee.

Lo scopo primario di LIBER è quello di supportare le biblioteche di ricerca europee con la creazione di una rete funzionale lungo i confini nazionali al fine di assicurare la conservazione del patrimonio culturale europeo, migliorare l'accesso alle collezioni delle biblioteche di ricerca europee ed offrire servizi informativi più efficienti entro l'Europa.

<http://www.kb.dk/liber/>

<sup>4</sup> SPARC Europe è un'alleanza di biblioteche di ricerca europee, organizzazioni bibliotecarie e istituzioni di ricerca che supporta forme competitive di editoria scientifica per i periodici. LIBER funge da ombrello organizzativo per SPARC Europe, la quale facilita la competizione nel mercato dei periodici scientifici europei e introduce e sostiene iniziative ritagliate sia sulla ricerca sia sulla comunità bibliotecaria in Europa.

<http://www.sparceurope.org/>

sponsorizzazione di *ESF European Science Foundation*<sup>6</sup>, *JISC Joint Information Systems Committee*<sup>7</sup>, e *OSI Open Society Institute*<sup>8</sup>.

I partecipanti alla conferenza sono stati 136, di cui otto italiani, molti provenienti dai settori accademici e di ricerca, dal mondo dell'editoria scientifica, alcuni presenti in qualità di membri degli enti organizzatori e/o sponsorizzatori, bibliotecari, tecnici e quanti interessati a discorsi sulla comunicazione scientifica open access.

Doveroso uno sguardo alla composizione del comitato organizzatore che ha avuto il merito di confezionare un incontro di indubbio successo sul piano scientifico dei contenuti, anche in relazione alla struttura delle sessioni suddivise per argomenti tecnici e casi studio di notevole interesse.

Nel comitato organizzatore spiccano infatti nomi tra i più importanti della biblioteconomia internazionale, molti di provenienza dalla taskforce di OAI, da OSI e membri LIBER: Chris Bailey bibliotecaria all'University Glasgow (CURL), Lars Bjørnshauge, direttore del sistema bibliotecario Università di Lund, Alison Buckholtz di SPARC, Raf Dekeyser bibliotecario all'Università di Leuven, e chairman di LIBER Access Division, Fred Friend per SCOUNL, Melissa Hagemann di OSI, Thomas Krichel, professore di biblioteconomia alla Palmer School of Library and Information Science, Università di Long Island a New York e fondatore di RePec, Corrado Pettenati, direttore della biblioteca del CERN, Christopher Pressler, capo della collezione d'arte alla biblioteca centrale dell'Università di Londra, Bas Savenije, bibliotecario all'Università di Utrecht, Ronald Schmidt, bibliotecario alla Hochschulbibliothekszenrum Nordrhein-Westfalen di Colonia, Jean-Philippe Schmitt,

---

<sup>5</sup> CERN European Organization for Nuclear Research, è il più grande centro del mondo per la fisica delle particelle. La biblioteca, assieme agli archivi scientifici del CERN sono incardinati entro la Divisione ETT-SI Scientific Information Service. La biblioteca del CERN oltre ad acquisire e gestire fonti informative in tutti i campi di rilevanza per l'organizzazione, rende accessibili tali risorse a tutta la comunità mondiale dei fisici delle particelle. In tale ottica di accesso esteso la biblioteca del CERN ha acquisito notevole esperienza nell'area della gestione e creazione di archivi di e-prints.

<http://library.cern.ch/>

<sup>6</sup> ESF European Science Foundation, agisce come catalizzatore per lo sviluppo della scienza, promuovendo un livello di alta qualità scientifica in Europa, consolidando, pianificando e implementando iniziative pan-europee.

<http://www.esf.org/>

<sup>7</sup> JISC Joint Information Systems Committee è una commissione strategica consultiva britannica, in partnership con il Consiglio d'Europa, che opera e nell'interesse degli istituti accademici ed educativi in merito a progetti specifici. JISC promuove e finanzia le applicazioni innovative e l'uso dei sistemi informativi e della tecnologia dell'informazione nei settori accademici e di alta educazione. Il suo ruolo centrale è assicurare un costante aggiornamento delle nuove tecnologie e dei metodi sull'analisi dei costi entro un'infrastruttura di rete che offra servizi informativi e materiali di alta qualità. Entro l'architettura del programma FAIR (Focus on Access to Institutional Resources) JISC sta finanziando 14 progetti, i quali comprendono una cooperazione con oltre 50 istituzioni ove sono coinvolti università, biblioteche, gallerie d'arte, collegi, musei ed enti commerciali.

FAIR si ispira all'iniziativa Open Archives dove le risorse digitali possono essere condivise tra le organizzazioni sulla base di un semplice meccanismo che consente ai metadati, esposti dalle risorse stesse, di essere raccolti entro i servizi.

<http://www.jisc.ac.uk/>

<sup>8</sup> OSI Open Society Institute, parte della Fondazione Soros, è una fondazione operativa e privata con sede a Budapest la quale supporta finanziariamente lo sviluppo e l'implementazione di una gamma di programmi di tipo culturale, educativo, multimediale, nell'ambito della società civile, pubblica amministrazione, salute pubblica, diritti umani e diritti delle donne e riforme economiche, legali e sociali. OSI è finanziata dal finanziere e filantropo George Soros. Ha una seconda base a New York la quale assieme alla sede di Budapest offre supporto e assistenza ad oltre 50 organizzazioni in tutto il mondo, all'interno della rete OSI, nei paesi dell'Europa Centrale e dell'Est, ex-Unione Sovietica, Guatemala, Haiti, Mongolia, Africa de Sud e Occidentale.

<http://www.osi.hu/>

responsabile del Groupe de compétences bibliothèques, RERO di Martigny, Herbert Van de Sompel, di LANL, Los Alamos National Laboratory, del comitato esecutivo OAI.

Grazie all'efficientissima organizzazione della biblioteca del CERN tutte le presentazioni sono state poste in linea fin da subito in vari formati (PDF, PPT, ...) e nelle versioni video e video sincronizzato con le slide<sup>9</sup>.

Per l'occasione è stata compilata da Sarah Farad un'utilissima bibliografia internazionale<sup>10</sup> sulla tematica dell'auto-archiviazione comprendente oltre un centinaio di riferimenti con link, per il periodo da gennaio a settembre 2002. Questo dato ci dimostra come la tematica dell'auto-archiviazione sia un filone di ricerca alquanto prolifico nel settore LIS.

Obiettivo del workshop era quello di offrire strumenti guida a quanti (persone o istituzioni) sono interessati a soluzioni open access per la disseminazione dei contenuti di ricerca, obiettivo ben raggiunto soprattutto entro le sessioni *breakout* dove i partecipanti hanno messo a fuoco le problematiche incontrate nelle iniziative intraprese e dove si è dibattuto sui processi di pianificazione, sui problemi tecnici e gestionali che intervengono nell'implementazione di un archivio per il deposito di e-prints OAI compatibile.

Il focus della conferenza si è svolto principalmente attorno a due canali, apparentemente distinti, ma di fatto interagenti l'uno con l'altro: gli archivi istituzionali da una parte, numerosi e a livello politico con una forza strategicamente innovativa, gli archivi disciplinari dall'altra, più ampi, che rappresentano momenti di forte aggregazione delle varie comunità scientifiche. Entrambi i modelli trovano unitarietà entro un'architettura mista, dove ciascun OpenArchive è interoperabile attraverso il protocollo OAI per la raccolta dei metadati.

Entro questo framework hanno trovato posto sia gli aspetti tecnici, sia quelli organizzativi, equamente distribuiti e discussi sia dai relatori sia entro le sessioni *breakout*.

Infatti dopo la prima sessione introduttiva, le due successive hanno riguardato i *case studies* su archivi organizzati a livello disciplinare, e i *case studies* su archivi istituzionali, dove in entrambi i fronti si sono trattati gli aspetti sia tecnici sia organizzativi.

In apertura **Herbert Van de Sompel**, giunto direttamente dalla conferenza internazionale sul Dublin Core 2002 che si teneva a Firenze più o meno (purtroppo) negli stessi giorni, ha illustrato la nuova versione del Protocollo OAI per la raccolta dei metadati, noto come OAI-PMH 2.0 OAI Protocol for Metadata Harvesting<sup>11</sup>, ora operativo dopo diciotto mesi di test sulla versione 1.1. OAI-PMH consiste di sei "verbi", tre dei quali rivelano la caratteristica dell'archivio (ListMetadataFormats, ListSets, and Identify) e tre l'estrazione dei metadati dall'archivio (GetRecord, ListRecords, ListIdentifiers).

Le novità hanno riguardato la chiara distinzione tra protocollo e periferia e la chiara separazione tra OAI-PMH e http, una miglior definizione di harvester, deposito, documento, identificatore unico, record, insieme, raccolta selettiva, utilizzo del *must* e *must not* ... come in RFC2119. Con la sua usuale verve e chiarezza espositiva Van de Sompel ha tracciato le linee fondamentali del movimento OAI delineando scenari dove solide conoscenze tecniche nella creazione e gestione di depositi OAI compatibili possono condurre ad assicurare il successo nella riforma dei processi di comunicazione scientifica.

---

<sup>9</sup> <http://doc.cern.ch/age?a02333>

<sup>10</sup> [http://documents.cern.ch/cgi-](http://documents.cern.ch/cgi-bin/txt2pre.sh?file=/archive/electronic/other/agenda/a02333/a02333/bibliography/Aself-archivingthematicbibliography.htm&style=cds)

[bin/txt2pre.sh?file=/archive/electronic/other/agenda/a02333/a02333/bibliography/Aself-archivingthematicbibliography.htm&style=cds](http://documents.cern.ch/cgi-bin/txt2pre.sh?file=/archive/electronic/other/agenda/a02333/a02333/bibliography/Aself-archivingthematicbibliography.htm&style=cds)

<sup>11</sup> <http://www.openarchives.org/OAI/openarchivesprotocol.html>

Entro la sessione che ha trattato il modello di deposito disciplinare tre sono state le relazioni. **Elizabeth Cherhal** ci ha parlato dell'esperienza francese Cellule MathDoc (Cellule de Coordination Documentaire Nationale pour les Mathématiques)<sup>12</sup> entro il CCSD (Centre pour la Communication Scientifique Directe). CellMathDoc nasce nel 1995 in seno al CNRS e al ministero della ricerca francese con lo scopo preciso di documentare a livello nazionale tutta la produzione intellettuale matematica, in cooperazione con le biblioteche e gli istituti matematici di Francia. CellMathDoc coopera con EMS European Mathematical Society e con IMU International Mathematical Union e in tale ambiente si è proposto come sviluppatore di software come per esempio lo strumento di ricerca basato su EDBM (European Database Manager for Mathematics) kernel, su server UNIX, per la ricerca entro Zentralblatt-MATH Database, dotato di numerose funzionalità tra cui la possibilità di una agevole navigazione su autori, codici di classificazione MSC<sup>13</sup>, citazioni, e recensioni.

**John Ober** ha illustrato il modello di OAI della California Digital Library, fondata nel 1997, da dieci campus universitari della California in risposta alla crisi della comunicazione scientifica. Entro tale modello è stato creato un sistema sperimentale per la gestione delle informazioni digitali ad uso della comunità scientifica, su base disciplinare.

The eScholarship Repository<sup>14</sup> è nato nel 2002 come archivio disciplinare per le scienze sociali e umane per l'archiviazione e disseminazione dei risultati di ricerca e working paper. E' costruito in partnership con la Berkeley Electronic Press (bepress), utilizzando strumenti per la sottomissione e relativa gestione dei processi attraverso un'interfaccia web tramite il sistema bepress EdiKit. Le configurazioni di EdiKit consentono di gestire i processi di peer-reviews entro l'archivio. Il deposito rappresenta una componente importante entro il CDL's eScholarship program, la cui missione è facilitare e supportare le innovazioni entro il circuito della comunicazione scientifica. CDL ha due tipi di accordi con bepress: uno per la licenza del software EdiKit per la gestione di periodici elettronici, e-book e working paper depositati dell'archivio, l'altro per lo sviluppo in cooperazione al fine di futuri miglioramenti.

**Susanne Dobratz** ha esposto le attività ed esperienze di Open Archives Forum<sup>15</sup>, il quale fornisce supporto a livello europeo per la disseminazione dell'informazione correlata agli Open Archive in Europa. L'obiettivo di OA Forum è quello di costruire una comunità di interesse europea utile allo scambio informativo su questioni tecniche e organizzative in merito all'implementazione di e-server OAI compatibili. OA-Forum fa parte del 5. programma quadro europeo, programma IST (Information Societies Technology). Tra i partner anche l'Italia con il CRN di Pisa. In tale assetto europeo mancano esperienze italiane, citate da Dobratz solo di volata, in quanto l'intensa attività e lo sforzo che si stanno compiendo nei nostri atenei (Padova e Bologna in particolare e al CNR Consiglio Nazionale delle Ricerche) in questi mesi, non sono noti a livello formale entro OAI. Sarà necessario fare

---

<sup>12</sup> <http://www-mathdoc.ujf-grenoble.fr/>

<sup>13</sup> Sempre nell'ambito degli strumenti software e di classificazione MSC, tra le righe si cita la versione della classificazione internazionale per la matematica, con interfaccia multilingue MSC2000 in francese, inglese, italiano: *Expérimental: navigation dans MSC, avec interrogation de banques de données Cellule de Coordination Documentaire Nationale pour les Mathématiques*  
<http://www-mathdoc.ujf-grenoble.fr/MS2000/db.html>

Idati per la versione italiana sono stati da noi (Università di Padova) messi a disposizione per il sito francese [www.math.unipd.it/~biblio/math/](http://www.math.unipd.it/~biblio/math/).

<sup>14</sup> <http://repositories.cdlib.org>

<sup>15</sup> <http://www.oaforum.org/index.php>

il salto di qualità e, per quanto ci riguarda, procedere con la registrazione in OAI degli Open Archive presenti nelle nostre università non appena operativi.

La sessione successiva ha riguardato i depositi istituzionali, in particolare **Jörgen Eriksson** dell'Università di Lund ha toccato gli aspetti pregnanti che intervengono nella costruzione di un archivio di tipo istituzionale, le difficoltà che si incontrano e le necessità di una crescita di consapevolezza da parte degli utenti. Un OA fatica a decollare e quindi è necessario pianificare ogni mossa al fine di evitare che esso rimanga un contenitore vuoto.

**Stephen Pinfield** ha descritto lo stato dell'arte degli Open Archive nelle università britanniche, descrivendo il progetto SHERPA Securing a Hybrid Environment for Research Preservation and Access<sup>16</sup>, finanziato dal JISC e da CURL Consortium of University Research Libraries e ospitato dall'Università di Nottingham. Scopo del progetto è investigare sui processi IPR (Intellectual Property Rights) sui processi di controllo di qualità, e altri aspetti chiave gestionali e culturali. Per gli aspetti tecnici SHERPA si occuperà di indagare sull'interoperabilità tra archivi e sulla conservazione digitale degli e-prints.

Pinfield ha segnalato l'importanza del progetto Romeo (*Rights Metadata for Open archiving*)<sup>17</sup>, più volte citato nell'ambito della conferenza anche da altri relatori, finanziato dal JISC (Joint Information Systems Committee) per la durata di un anno (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. Il progetto entro un quadro OAI tiene conto delle questioni correlate alla raccolta dei metadati dai data providers entro i service providers. Romeo coordinato da Charles Oppenheim dell'Università di Loughborough, si occuperà 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 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.

Un altro progetto ripreso da Pinfield è TARDIS (*Targetting Academic Research for Deposit and Disclosure*), coordinato dall'Università di Southampton, il quale investiga sul superamento delle barriere tecniche culturali ed accademiche nei depositi istituzionali. TRADIS svilupperà inoltre modelli di lavoro per i depositi multi-disciplinari.

**Gerhard Beier** del Max Planck Institute ha illustrato l'esperienza dell'archivio istituzionale eDoc<sup>18</sup> che serve differenti comunità afferenti a ottanta istituti.

eDoc utilizza strumenti come BibTex, Reference Manager, Endnote per l'import dei dati e implementa MPG-SFX per il reference linking sensitivo ai fini dell'integrazione e recupero dei documenti a testo pieno.

---

<sup>16</sup> <http://www.nottingham.ac.uk/isd/newsletter/inform-online-7.2/sherpa.htm>

<sup>17</sup> Informazioni sul Progetto Romeo (Rights Metadata for Open archiving)

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

<sup>18</sup> <http://edoc.mpg.de/>

L'università di Lubiana ha portato il progetto SciX<sup>19</sup> esposto da **Ziga Turk** il quale ha focalizzato il suo intervento sui vantaggi di un modello di editoria scientifica che abbassa le barriere tecniche, economiche e sociali. SciX è un progetto europeo con termine nel 2004 al quale partecipano sette partner europei. Lo scopo di SciX è quello di offrire ai ricercatori gli strumenti per creare archivi digitali al posto delle pagine web dove attualmente vengono posti di solito i lavori, creando in questo modo una comunità on-line di autori e lettori in cui i lavori scientifici sono disseminati anche ad altre comunità correlate ad altri settori di mercato.

In questa sessione è risultata di grande interesse la relazione di **Smith MacKenzie** sulla piattaforma *DSpace*<sup>20</sup>, prodotto sviluppato nel 2000 entro un progetto congiunto dalle biblioteche del MIT (Massachusetts Institute Technology) e dalla Hewlett-Packard Company per la costruzione di un repository a lungo termine, stabile e sostenibile, che consenta la cattura, la conservazione e la disseminazione della ricerca e dei materiali didattici, generati lungo lo spettro delle discipline al MIT, consentendo accesso remoto via web da qualsiasi punto della rete.

Poiché molte delle esperienze attuali, soprattutto per quanto concerne gli archivi di tipo istituzionale, utilizzano il software GNU EPrints (di cui si parlerà in seguito), l'annuncio di questa nuovo prodotto ha incuriosito molti dei presenti, in particolare perché, a quanto è stato detto, *DSpace* offre i vantaggi di una distribuzione e conservazione digitale per una vasta gamma di formati inclusi testi, audio, video, immagini, programmi e datasets. Tutti i materiali possono essere soggetti a restrizioni d'accesso ai fini del copyright elettronico e quindi è possibile gestire gli accessi a fasce di utenza differenziata (studenti, ricercatori, staff). Come illustrato da MacKenzie le due configurazioni primarie del sistema *DSpace* si basano sui due presupposti distinti: il primo si riferisce all'aspettativa posta sugli autori che, attraverso l'auto-archiviazione dei propri lavori, forniscono direttamente i metadati alimentando la base di dati, come nel sistema EPrints Harnad/Gutteridge del resto. Il secondo è basato sulla considerazione che l'archivio consente di gestire una grande varietà di formati e tipi di oggetti digitali. Infatti da una parte *DSpace* nasce con lo scopo di fornire record sostenibili riferibili alla produzione intellettuale dei ricercatori afferenti alle facoltà del MIT, dall'altra ha il preciso obiettivo di offrire un valido supporto alla gestione della produzione multimediale didattica, entro un ambiente configurabile come e-learning.

Il software *DSpace* sarà reso disponibile pubblicamente a partire dal 4 novembre 2002, attraverso licenza BSD OpenSource, la quale consente ad altre istituzioni di utilizzare il software *DSpace* per creare propri archivi per il deposito.

Il secondo giorno si è suddiviso in due parti, nella mattinata si sono esplorati gli argomenti tecnici, il pomeriggio si è svolto all'insegna del motto "guadagnando l'indipendenza".

La sessione mattutina coordinata da **Corrado Pettenati** ha passato in rassegna argomenti di notevole portata, quali la raccolta di metadati, le infrastrutture e le architetture entro un modello distribuito su vasta scala, i service provider, i motori di ricerca, l'archiviazione a lungo termine, gli OpenArchive entro biblioteche digitali.

**Femke Markus** di Elsevier Science, tra interventi del pubblico piuttosto polemici, è riuscita a illustrare il piano di sviluppo, il modello economico e le motivazioni che hanno condotto alla

---

<sup>19</sup> <http://www.scix.net/>

<sup>20</sup> <http://www.dspace.org/>

creazione del motore di ricerca Scirus<sup>21</sup>, il quale ricerca oltretutto tra le risorse Elsevier, entro alcuni database free access come PubMed e in alcuni Open Archive tra cui arXiv di LANL.

Il tema dell'archiviazione a lungo termine è stato trattato da **Christopher Pressler** dell'Università di Londra, il quale ci ha offerto una panoramica focalizzata sulle attività JISC le quali riguardano numerosi aspetti tra cui la degradazione dei media, l'obsolescenza tecnologica, l'autenticità, la proprietà intellettuale, i costi di deposito correlati al ciclo di vita delle risorse.

I media degradano e le informazioni sono facilmente deperibili se non si adottano accortezze adeguate, i formati subiscono un processo di rapida obsolescenza e sono accessibili attraverso hardware e software che mutano nel tempo. Tutti i record digitali in pratica devono essere spostati da un sistema all'altro attraverso l'evolversi delle nuove tecnologie e questo comporta una serie di compromessi su azioni da condurre e sui costi da sostenere.

A supporto di queste tematiche il JISC, nell'ambito dei suoi progetti, offre una serie di servizi, standard di conservazione per il deposito, certificazioni e strumenti utili ad una vasta gamma di centri distribuiti che si occupano di dati digitali o che effettuano servizi di conservazione, tra cui lo sviluppo di un deposito di software, documentazione e strumenti sulla conservazione digitale. Per quanto riguarda la proprietà intellettuale Pressler ha ricordato che la legge britannica non prevede nell'ambito del fair dealing o dei privilegi alle biblioteche un diritto automatico di riproduzione archiviazione a scopo di conservazione.

**Michael L. Nelson** ci ha offerto una ricca panoramica storica e sulle prospettive future dei service provider entro il quadro OAI-PMH. Nelson ha esordito tracciando, attraverso l'analisi dei precedenti meeting OAI; lo sviluppo di alcune tematiche a livello storico. Se nel primo meeting OAI Open Day di Washington nel gennaio 2001 quattro furono gli item sulla definizione di protocolli, sviluppo di strumenti, al presente meeting ginevrino solo uno ha riguardato tale tematica. L'adozione di nuovi servizi e di service provider in particolare ha subito un notevole incremento rispetto al passato, come pure la tematica correlata agli aspetti politici e socio-economici. La chiave per comprendere la filosofia di OAI è capire la separazione di responsabilità dei data provider (DP) e dei service provider (SP). In pratica, seppur distinti, sia un DP sia un SP possono risiedere nella stessa entità. Un data provider è un deposito o archivio o anche una semplice collezione di metadati (record) i quali possono come no, puntare al corrispondente documento a testo pieno. Un service provider offre un servizio a valore aggiunto, per esempio attraverso il browsing o la ricerca sui metadati estratti da uno o più data provider (archivi). Nelson ci ha offerto una carrellata degli attuali SP disponibili: *NTRS* della NASA, il recente *Arc A Cross Archive Search Service* (nato sul precedente modello *UPS Universal Preprint Service*<sup>22</sup>), *NCSTRL Networked Computer Sciences Technical Reference Library*, *Archon* (una versione Arc per i fisici), *Torii* sviluppato dalla SISSA di Trieste, una specie di portale per i fisici su desktop personalizzato e con possibilità di accesso via WAP, *Icite, my.OAI, Cyclades, citebase* per il reference linking basato sul software Eprints, gli harvester *OAIster* e *Public Knowledge Project* e *Perseus*. E' chiaro, ci dice Nelson, che i service provider stanno proliferando a causa del bias verso i data provider

---

<sup>21</sup> <http://www.scirus.com/>

<sup>22</sup> Herbert Van de Sompel, Thomas Krichel, Patrick Hochstenbach; Victor M. Lyapunov, Kurt Maly, Mohammad Zubair, Mohamed Kholief, and Xiaoming Liu, Heath O'Connell  
*The UPS Prototype: An Experimental End-User Service across E-Print Archives*  
<http://www.dlib.org/dlib/february00/vandesompel-ups/02vandesompel-ups.html>

entro il protocollo. Nelson ha anche spiegato i motivi per cui OAI-PMH non è importante, sottolineando che l'utilizzo del protocollo non assicura comunque un buon service provider. Ciascun SP è libero di definire i propri servizi, la modalità di presentazione e personalizzazione delle interfacce all'utenza. Questi servizi potrebbero perciò divenire assai competitivi.

Di sicuro interesse l'esperienza del Caltech<sup>23</sup> California Institute of Technology, denominata CODA Collection of Open Digital Archives, illustrata da **Eric Van de Velde**, modello a cui guardare sia per l'efficienza organizzativa sia per le scelte tecniche operate in merito alla raccolta di metadati.

Nel modello di Caltech intervengono 285 corsi/facoltà, 900 studenti undergraduate, 1100 studenti graduate e 1000 ricercatori. Il sistema bibliotecario risulta composto da 55 unità FTE, 14 bibliotecari, 6 tecnici, 500.000 volumi, 3556 abbonamenti a periodici a stampa e 2116 abbonamenti a periodici elettronici, 88 banche dati online, 160 postazioni, 100 per il pubblico, 4 grosse biblioteche d'area + 2 locali, un budget annuale di \$ 6.370.000.

La scelta CODA si è basata sull'utilizzo del software GNU EPrints e, per le tesi, sul software NDLTD. CODA si è concretizzato con la creazione di numerosi archivi. Attualmente i documenti presenti sono oltre 1100. I vantaggi di una frammentazione del deposito in più archivi ha consentito un controllo amministrativo idoneo a ciascuna collezione, per esempio la revisione sui formati e/o sui contenuti, l'assegnazione dei numeri di report, la gestione degli accessi, la politica di raccolta dei metadati e la politica stessa sui metadati da adottare per ogni singola collezione.

La differenziazione del deposito in più archivi ha comportato anche alcuni svantaggi, non ultimo un appesantimento del lavoro dovuto alla gestione e al mantenimento degli stessi.

La distribuzione dei repositories CODA risulta composta da: un archivio per le tesi noto come NDLTD Networked Digital Library of Theses and Dissertations<sup>24</sup> dove avviene la diretta sottomissione delle tesi da parte degli autori a partire dal luglio 2002 (immissione delle tesi scansionate dallo staff a partire dal 1998) che contiene 294 tesi liberamente disponibili e 49 ad accesso ristretto, un archivio per i rapporti tecnici che contiene circa 60 papers auto-archiviati, per la maggior parte provenienti dal settore scientifico, un archivio per gli atti delle conferenze per 368 documenti disponibili, un archivio per i report di ambito LIS con 15 lavori, uno specializzato per il progetto "Earthquake Engineering", un archivio con un solo volume (per il momento) e una serie di altri archivi per collezioni non di ricerca e altri in corso di implementazione sui settori dell'aeronautica, meccanica dei fluidi e dei solidi, controllo e sistemi dinamici, ingegneria meccanica, ...

L'insieme di questi repositories in parte disciplinare, in parte per collezione, crea un unico insieme di repository istituzionale composto da tanti data provider entro un quadro di service provider.

Notevoli le attività sul controllo dei metadati, l'adozione di formati e la formazione sulle tematiche relative al copyright, in particolare per l'archivio tesi.

Molto lavoro è stato fatto per la predisposizione delle politiche di deposito e di copyright, considerando che la partecipazione era stabilita su base volontaria, e considerando che doveva essere attuato un controllo su chi deposita cosa, attraverso un processo di approvazione interno. Le politiche hanno dovuto definire questioni fondamentali quali per esempio rendere o meno permanente la sottomissione dei lavori, concedere un permesso di distribuzione non esclusivo (diritti non esclusivi) e stabilire una disponibilità piena per lo scambio di metadati.

---

<sup>23</sup> <http://coda.caltech.edu>

<sup>24</sup> <http://www.ndltd.org/>

Sono state stilate delle licenze di utilizzo per gli autori e per gli utenti che determinano con chiarezza la questione del copyright elettronico.

Tale modello organizzativo è un utile punto di riferimento che andrebbe calato nelle singole realtà nazionali che possono ovviamente presentare delle divergenze in termini di normative sulla proprietà intellettuale, ma anche su aspetti più prettamente correlati alla disponibilità di risorse (finanziarie e di personale).

Altro punto di estremo interesse nell'esperienza di Caltech è l'applicazione di un risolutore<sup>25</sup> per tutti i documenti, una sorta di URL permanente ai documenti depositati entro il server, più economico del DOI, che tiene traccia delle localizzazioni indipendentemente da possibili futuri cambiamenti di software. PURR Persistent URL Resource Resolver si basa su un meccanismo che utilizza l'identificatore con sintassi stabilita in una determinata forma.

Con ritmo brillante e metodologia rigorosa **Thomas Krichel** ha condotto una presentazione da economista, in perfetto stile bibliotecario. Le slide dai titoli assai eloquenti, difficilmente traducibili in italiano, ci hanno guidato entro gli aspetti strategici, economici e sociali, che caratterizzano la costruzione di reti di archivi su vasta scala, in un gioco di parole chiave da afferrare al volo: "cercare alleati" ("searching for allies") "il mondo è tuo, prendilo" ("the world is yours to snarf"), "a lungo termine siamo tutti morti" ("in the long run we are all dead"), "la teoria dell'aggregazione valutativa" ("theory of evaluative aggregation"), "coinvolgere il mondo accademico" ("getting sholars involved"), "autorità!, disse" ("authority, he said"), "status quo", "status in spe", "grido di battaglia" ("rallying cry"),...

La chiave del successo, ci dice Krichel sta nell'ottenere che altri si aggregino e cooperino attivamente e fare in modo che altri facciano la propria parte. Fondamentalmente si tratta di un problema economico e le soluzioni stanno nelle potenzialità del lavoro volontario distribuito e nella creazione di alleanze. Per guardare avanti è necessario applicare le teorie specifiche dell'architettura economica, ma soprattutto risulta strategico cercare alleanze. Krichel come lo ha definito Billy Arms è noto per aver fatto molto con risorse assai limitate. Partito nel 1993 senza dati, senza nemmeno un computer entro cui inserire i dati, per Krichel le alleanze sono state cruciali nella creazione di quella che è considerata la prima e più grande rete del mondo di OpenArchive a modello distribuito e organizzata su base disciplinare. Infatti nel 1997 fonda RePEc<sup>26</sup>, una biblioteca digitale per le discipline economiche, la quale si basa su un'architettura che implementa data provider versus service provider su un modello tecnicamente robusto seppur non raffinato rispetto a sistemi interattivi online, basato sulla raccolta di file statici e sull'utilizzo di template. E' necessario avere un quadro di ciò che si vuole raggiungere prima di decidere come raggiungerlo in quanto le trasformazioni sociali sono più importanti di qualsiasi configurazione tecnologica.

L'editoria scientifica è più un servizio orientato all'autore piuttosto che al lettore, quindi sarà necessario tener conto della teoria dell'aggregazione valutativa. Per questo dobbiamo impossessarci della comunicazione scientifica come processo organico entro una comunità.

La comunità scientifica deve essere coinvolta, per questa ragione risultano più opportuni gli approcci basati su disciplina i quali non sono affatto in contraddizione con gli archivi di tipo istituzionale. Krichel ha puntato molto sulla necessità di ottenere consenso dai ricercatori, ma

---

<sup>25</sup> Uno script estrae i dati da un archivio EPrints e li importa per un uso immediato dentro al sistema PURR. Lo script sincronizza l'informazione dal database EPrints con il database PURR. Ogni archivio deve avere un proprio database PURR. Componente chiave è un programmino in Perl scaricabile liberamente. Per l'implementazione serve: ambiente Linux, MySQL, Apache, Perl, mod-perl, e DBI.  
<http://resolver.library.caltech.edu/CaltechBOOK:1995.001>

<sup>26</sup> RePEc Research Papers in Economics  
<http://repec.org/>

per questo abbiamo bisogno di presentare loro modi per misurare la qualità di quello che stanno facendo. Andranno quindi raccolti e aggregatori dati per la valutazione in merito a visualizzazioni dell'abstract, scarico del documento, analisi delle citazioni, ... I "dati valutativi" faranno in modo che gli scienziati ci vengano dietro. Abbiamo bisogno di strutturare gli archivi, le autorità e i servizi. Abbiamo bisogno di metadati più specifici, e di metadati più relazionali, di aggregati di dati. Entro gli OpenArchive abbiamo bisogno di insiemi di dati secondari aperti: persone, istituzioni, documenti, collezioni di documenti, ma per questo ancor prima abbiamo bisogno di dataset relazionali. Krichel ha focalizzato sull'importanza dei gruppi di autorità . Un modello di autorità è fornito da RePEC, altri emergenti sono rclis<sup>27</sup> e Physnet<sup>28</sup>.

Con il suo linguaggio molto slang Krichel conclude dicendo "Se un branco qualsiasi di volontari può offrire un completo sistema operativo in Internet allora non c'è ragione per dire che un gruppo di bibliotecari non possa offrire un catalogo completo di dati scientifici in Internet".

**Carl Lagoze** della Cornell University ci ha parlato della National Science Digital Library NSDL<sup>29</sup>, biblioteca digitale su larga scala sorta nel 1996 che costruisce tutti i suoi servizi attraverso strati di conoscenza su una varietà di risorse per una varietà di utenti. La filosofia portante di NSDL è basata sulla possibilità di costruire una biblioteca digitale davvero grande con un piccolo staff a disposizione, ma per far questo occorre che tutti gli aspetti biblioteconomici siano pianificati con una certa scalabilità programmata e accettando alcuni compromessi inevitabili. Alcuni *mantra* tecnici, come lui li definisce sono: aggregare piuttosto che collezionare di modo che il team non debba essere costretto a gestire direttamente le collezioni, accogliere diversi modelli di partecipazione entro uno spettro di interoperabilità con interfacce aperte e standard che consentano plug-in per servizi a valore aggiunto, ritagliare la presentazione dei contenuti e la natura dei servizi sulla base delle necessità degli utenti entro un'unica biblioteca digitale munita però di differenti porte di accesso (portali), aprire un toolkit di software e servizi utili alla costruzione di biblioteche digitali. Lagoze ha parlato di un metadata repository open access entro NSDL di supporto ai costruttori di servizi via OAI-PMH, poggiante su una strategia di metadati che fornisca adeguati crosswalks dagli otto formati standard previsti al Dublin Core. SiMBaS Simple Metadata-Based Services è il modello creato dal team NSDL per lo sviluppo e implementazione di servizi bibliotecari che dovrebbero essere basati sulla semplice estensione di records di metadati standard.

A proposito dello spettro di interoperabilità Lagoze ha illustrato i tre livelli di approccio in cui i metadati possono assumere varie forme, il livello di federazione, quello di harvesting e il terzo detto di gathering. Il livello federativo prevede stretto uso di standard quali MARC o z39.50, il livello harvesting, tipico degli ambienti OpenArchive prevede protocolli semplici per la raccolta dei metadati esposti dalle biblioteche digitali, mentre il livello di gathering è quello dei motori di ricerca o webcrawler dove le di biblioteche digitali non cooperano e quindi i servizi informativi si rivolgono all'esterno per la ricerca delle informazioni.

---

<sup>27</sup> research in computing, library and information science

<http://rclis.org>

<sup>28</sup> PhysNet - the worldwide Network of Physics Departments and Documents

<http://www.physics-network.org/PhysNet/>

<sup>29</sup> <http://about.nsdlib.org/>

La conferenza ha ruotato attorno alle tre iniziative principali, Open Archives Initiative, Budapest Open Access Initiative<sup>30</sup>, e SPARC's Gaining Independence, tutte strettamente allacciate entro un quadro di una comunità scientifica in grado di controllare autonomamente le proprie produzioni intellettuali al fine di servire meglio la mission delle nostre istituzioni accademiche e di ricerca.

Il workshop infatti mutua il nome "Guadagnando l'indipendenza" proprio dall'iniziativa SPARC<sup>31</sup> illustrata da **Alison Buckholtz** la quale offre un approccio pratico alla pianificazione e implementazione di esperienze di pubblicazione elettronica che, se da una parte possono rivelarsi assai competitive, dall'altra possono risultare rischiose. Come si è visto da alcune relazioni successive sul tema, la creazione di un periodico elettronico alternativo ai tradizionali modelli necessita di una preventiva costruzione di una rete di supporti e di una chiara strategia comunicativa alla base, entrambe poggianti su modelli di depositi digitali che veicolino una pianificazione progettuale in termini economici a lungo termine.

E' lungo il modello OAI che la piena indipendenza trova spazio al fine di garantire anche un pieno controllo sulla disseminazione dei contenuti, ma prima ancora sulla loro validazione in termini di qualità.

OAI come offerta di strumenti tecnologici che consentono di guadagnare l'indipendenza, si pone come passaggio obbligato nell'utilizzo esteso di protocolli, standard e software per la costruzione di archivi o depositi. BOAI si pone invece come momento esteso di sensibilizzazione degli autori (ricercatori, scienziati), verso una presa di coscienza collettiva che conduca ad un'auto-archiviazione consapevole della propria produzione intellettuale.

L'iniziativa di Budapest presentata da **Frederick J. Friend** (University College London) si distingue da OAI giacché il suo scopo è di fare pressione sugli autori al fine di ottenere un consenso generalizzato. Per BOAI ciò che conta è il consenso dell'autore, non soltanto le necessità o desideri 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.

In tale simbiosi di iniziative, i bibliotecari sono il fulcro operativo, da una parte gestori dell'informazione posta in luoghi dai nuovi orizzonti (gli Open Archive) dall'altra motori propulsori in un processo formativo ad ampio spettro su tematiche calde, quali appunto l'auto-archiviazione, le politiche di sottomissione dei dati, le politiche di copyright compresa la questione correlata al *metadata copyright*.

Questi aspetti, strettamente correlati, giocano un ruolo fondamentale entro il processo in trasformazione della comunicazione scientifica.

Nella politica dei costi delle riviste scientifiche, il monitoraggio degli input, gli aspetti finanziari, di copyright e la rete di relazioni con altri atenei e comitati editoriali, sono punti fondamentali di un processo in trasformazione, come si è visto nella relazione di **Ulf Rehmann** dell'Università di Bielefeld, il quale ha illustrato i sette anni di esperienza con la rivista *Documenta mathematica*<sup>32</sup>, gestita entro SPARC. *Documenta mathematica* è il tipico

---

<sup>30</sup> <http://www.soros.org/openaccess/>

<sup>31</sup> Gaining Independence: A Manual a cura di SPARC Scholarly Publishing and Academic Resources Coalition  
<http://www.arl.org/sparc/GI/>

<sup>32</sup> <http://www.math.uiuc.edu/documenta/>

esempio di rivista elettronica gestita completamente entro la comunità dei parlanti, in questo caso i matematici<sup>33</sup>.

Da sottolineare nell'esposizione di Rehmann il citato rapporto dal titolo "*Math Journal Price Survey*"<sup>34</sup> allegato alla presentazione, basato sulla precedente indagine svolta da AMS per gli anni 1994-2001 (ora peraltro aggiornata al 2002)<sup>35</sup>. Il lavoro è estremamente interessante soprattutto come strumento di riferimento per la valutazione dei costi di abbonamento delle riviste matematiche nel tempo. Dal lavoro di comparazione di Rehmann emerge che per le riviste matematiche vi è stato un incremento annuo del 17% di media durante i passati sei anni. "*This is, the overall increase factor in that case is 1.176 = 2.57, e.g., prices have more than doubled within the last 6 years.*"

In un'ottica analoga si pone anche il progetto FIGARO Federated Network of European Academic Publishers<sup>36</sup>, esposto da **Bas Savenije** dell'Università di Utrecht, modello di pubblicazione elettronica entro un ambiente accademico correlato all'iniziativa olandese Roquade. FIGARO è basato su un modello economico a tre strati: back office, un circolo di front-office e una moltitudine di iniziative editoriali entro l'istituzione accademica. La piattaforma sarà operativa alla fine del 2004, nel frattempo la coesione di più partnership ha consentito il consolidamento di una pratica sulla formalizzazione di contratti ed accordi.

**Jean-Yves Le Meur** del CERN ha condotto una presentazione assai efficace sul prodotto CDSware (CERN Document Server Software)<sup>37</sup>. CDSware è un insieme di moduli software, in soluzione OAI compatibile, per la gestione di collezioni di dati, sviluppato dai tecnici del CERN che consente la ricerca incrociata su set di archivi differenti. CDSware è un software libero e viene rilasciato gratuitamente sotto licenza di tipo GPL GNU, interamente costruito con software libero: Apache, MySQL, PHP, Python, WML. Implementa funzioni sia di data provider sia di service provider ed è in grado di scambiare metadati tra repositories eterogenei.

Le principali configurazioni di CDSware sono: interfaccia configurabile come portale atta ad ospitare varie tipologie di collezioni, potente motore di ricerca con sintassi simile a Google, personalizzazioni utente, incluse *document basket* ed *email notification alert*, sottomissione elettronica dei documenti e caricamento di vari tipi e formati di documenti.

Consente di ricercare entro e-print server, dentro il database catalogafico (catalogo) o altro archivio di documenti su web. E' inoltre implementato il protocollo di harvesting per la raccolta dei metadati OAI (versione 2.0). Usa MARC21 come standard bibliografico.

I moduli principali che compongono CDSware sono: *WebSubmit*, l'OpenArchive integrato che consente la sottomissione dei dati, conversione automatica dei formati, generazione del numero automatico di report, funzioni multiple per la post-sottomissione. I due moduli *BibHarvest* e *BibConvert* permettono di eseguire importazioni massicce di record da fonti diverse, attraverso template per la descrizione delle fonti da cui eseguire il caricamento o per descrivere la trasformazione (modifica) delle fonti stesse. Il modulo *BibFormat* è adibito alla formattazione flessibile e connessione dei record (linking). La componente *WebSearch* supporta funzionalità di ricerca e meta-ricerca, formati diversi tra cui PostScript, PDF,

---

<sup>33</sup> chi scrive proviene da una esperienza ventennale di gestione informativa entro la comunità matematica dell'Università di Padova (Italy)

<sup>34</sup> [http://www.mathematik.uni-bielefeld.de/~rehmann/BIB/AMS/Price\\_per\\_Page.html](http://www.mathematik.uni-bielefeld.de/~rehmann/BIB/AMS/Price_per_Page.html)

<sup>35</sup> <http://www.ams.org/membership/journal-survey.html>

<sup>36</sup> <http://www.figaro-europe.net/>

<sup>37</sup> <http://cdsware.cern.ch>

Msword, MSeXcel, MSPowerPoint. Questo modulo può essere considerato come un OPAC evoluto o meglio un SuperOPAC che consente di ricercare attraverso funzionalità molto spinte sia nei metadati, sia nelle citazioni, sia nel full-text dei documenti. Le interfacce di ricerca sono numerose e personalizzabili e quindi WebSearch quale SearchEngine è in grado di ricercare in tutte le collezioni presenti, e non solo nel gestionale (catalogo). *BibWord* si occupa dell'indicizzazione dei dati in connessione con il modulo WebSearch, creando gli indici a partire dai metadati e dallo stesso full-text al fine di un recupero dell'informazione più appropriato alle query poste. *WebAccess* è il modulo che consente di trattare gerarchie complesse, come le viste sui soggetti per un browsing espanso a più livelli. *WebPerso* è adibito alla personalizzazione dell'accesso web e *BibData* alla modifica records (solo per lo staff dei bibliotecari)

**Chris Gutteridge** dell'Università di Southampton (UK), considerato da tutti il guru di EPrints, ci ha offerto una panoramica sul sistema GNU Eprints<sup>38</sup>, nella sua ultima versione 2.1.1, un particolare esemplare di data provider, OAI compatibile, pensato per il deposito o l'auto-archiviazione di documenti in formato elettronico, originariamente orientato, ma non limitato, a contenuti legati alla ricerca scientifica. Nato da un'idea di Stevan Harnad, di cui si è sentita la mancanza in aula, soprattutto in sede di dibattito, è stato sviluppato presso l'Electronics and Computer Science Department dell'Università di Southampton dallo stesso Christopher Gutteridge, con la collaborazione di Mike Jewell, mentre la progettazione e la realizzazione della versione 1.0 sono di Robert Tansley e risalgono al 2001. GNU EPrints è un software liberamente utilizzabile, modificabile e redistribuibile, assieme al codice sorgente, in base alla licenza GNU General Public License. E' costituito per la gran parte di script CGI in linguaggio Perl, operanti su server http Apache e utilizza il DBMS relazionale MySQL. Utilizza solo standard aperti. Non richiede hardware particolarmente potente, né sul lato server né sul lato client, su cui è sufficiente un browser web.

E' relativamente semplice da usare ma al contempo potente, per lo scopo per cui nasce, anche per il fatto che è compatibile con il protocollo di harvesting di OAI.

Il sistema è dotato di funzioni di memorizzazione e alert automatico via email per ricerche personalizzate, funzioni disponibili per tutti gli utenti registrati.

A fine ottobre dovrebbe uscire la versione 2.2. che conterrà una serie di ulteriori miglioramenti.

A differenza dell'anno precedente, tra i partecipanti vi è stato un numero maggiore di persone che, nell'ambito delle loro istituzioni o a differente livello organizzativo, avevano già una certa esperienza di sistemi per l'implementazione e gestione di archivi di deposito. Tale nucleo che ha potuto scambiarsi preziose informazioni durante i momenti di dibattito o nelle pause, in merito ai punti di criticità tecnico-organizzativi, è stato di aiuto a quei partecipanti che avevano nozioni di OA solo teoriche e quindi necessitavano di un supporto informativo di tipo pragmatico, essenziale ad una implementazione vera e propria.

All'interno degli otto panels delle sessioni *breakout* i partecipanti hanno potuto confrontarsi e discutere su alcune tematiche calde utili ad intraprendere percorsi comuni. Il gruppo italiano si è suddiviso entro panels differenti.

Nella prima sessione coordinata da **Thomas Krichel**, alla quale ho partecipato e sulla quale posso relazionare in modo più approfondito, si è discusso in merito all'implementazione di depositi (data provider) entro le università e i centri di ricerca. Quali sono gli ostacoli al progredire degli archivi scientifici, quali gli elementi base richiesti per mettere in piedi un

---

<sup>38</sup> <http://software.eprints.org/>

archivio, quali le strategie che le università utilizzano per rendere operativi i propri archivi in termini di interoperabilità con altri atenei.

In tale ottica sarà quanto mai necessario un approccio cooperativo piuttosto che competitivo tra gli atenei italiani. In particolare per quanto riguarda gli aspetti organizzativi, quali il numero di archivi da istituire, l'albero dei soggetti da utilizzare, sarà strategico per un successo italiano puntare su un coordinamento nazionale.

Tale coordinamento dovrebbe anche occuparsi delle configurazioni per l'uscita dei metadati OAI Dublin Core per un harvesting quanto più possibile formalizzato in termini di interoperabilità nella prospettiva di un service provider italiano, suddiviso per discipline, entro un modello distribuito di <n> data provider istituzionali accademici e di ricerca.

I processi che caratterizzano gli attuali circuiti della comunicazione scientifica sono percepiti da più parti come fortemente limitanti nei confronti di una diffusione appropriata e rapida dei contenuti, a causa di una gestione della proprietà intellettuale inadeguata, sempre più monopolio degli editori.

E' necessario educare gli utenti-autori, ma soprattutto lo staff dei bibliotecari, a vivere un rapporto con il copyright più sereno; la paura è nemica, quindi una sua comprensione approfondita aiuta a superare alcuni ostacoli certe volte solo apparenti e al contempo previene dal commettere violazioni facilmente evitabili.

E' anche necessario educare gli utenti al corretto uso degli strumenti per la multimedialità, ma soprattutto coinvolgerli nella politica della sottomissione dei lavori.

Altro problema fortemente sentito in alcuni gruppi di discussione è quello correlato all'acquisizione di tecnologia, non solo di risorse, ma di competenze. E' più che mai necessario convertire bibliotecari tradizionali in figure di bibliotecari tecnici che siano in grado di lavorare in ambiente Linux e di operare con applicazioni indispensabili alla gestione degli Open Archive.

Il secondo panel coordinato da **Fred Friend e Alison Buckholtz** ha riguardato gli aspetti e le problematiche che si instaurano nelle relazioni umane attorno ai server istituzionali. Problemi riguardanti da una parte l'approccio dell'autore come soggetto auto-depositante e quindi catapultato entro un modello nuovo in cui teme influenze ed effetti negativi sulla sua carriera, dall'altra i problemi dell'utente come soggetto che cerca e vuol recuperare lavori di qualità, spesso all'oscuro delle norme che regolano le attività correlate alla proprietà intellettuale (riproduzione, citazione, scarico dei lavori altrui). Di mezzo i problemi gestionali che riguardano i costi necessari (finanziari e di risorse) per instaurare processi di qualità alla base del successo.

La terza sessione in francese è stata coordinata da **Helene Bosc** e ha trattato i problemi tecnici e finanziari nello sviluppo e crescita di un OAI. Nella sessione coordinata da **Herbert Van de Sompel e Carl Lagoze**, molto tecnica, si è discusso del protocollo OAI, mentre in quella coordinata da **Van de Velde** si sono trattati i software OAI, quali sono i software per la costruzione e gestione di un archivio, quale è il più appropriato per ciascuna struttura, quali i requisiti necessari ad una sua implementazione e quali le configurazioni offerte da ciascun prodotto ad oggi disponibile.

**Susan Dobratz** ha condotto la sessione sull'archiviazione a lungo termine lungo il motto "guadagnando l'indipendenza con gli archivi di e-print OAI compatibili". Le linee guida lungo le quali si è mossa la discussione del suo gruppo hanno toccato le figure o attori che partecipano alle possibili attività entro un OA: i produttori di informazione, i revisori interni o comitato editoriale, gli operatori e gestori del deposito, gli utenti. I produttori informativi dovranno tener conto di alcune regole tra le quali l'utilizzo di formati adeguati, chiarezza di

note sul copyright, utilizzo di firma digitale. Un appropriato controllo di qualità sia formale che di contenuto dovrà essere applicato alle informazioni depositate. Questo sarà compito dei revisori o comitati editoriali i quali dovranno anche occuparsi della definizione dei canali di distribuzione entro il processo di validazione.

Gli operatori dovranno piuttosto garantire la sicurezza tecnica del sistema, l'adeguamento agli standard tecnici ai protocolli e ai modelli funzionali per l'interoperabilità, organizzando una viabilità delle informazioni entro un sistema sostenibile a livello finanziario dove ciascuno degli attori assume la piena responsabilità del proprio ruolo. Per quanto riguarda gli utenti, che in parte sono gli stessi autori, è necessario che divengano consapevoli dell'uso appropriato dei materiali di cui possono fruire.

Il gruppo coordinato da **Michael L. Nelson** ha individuato numerose aree su cui focalizzare i futuri sviluppi dei servizi entro l'architettura OAI. La creazione di VAA *Value Added Aggregator*, aggregatori di valore aggiunto, dovrebbe conformare molti dei servizi associati agli attuali tradizionali servizi offerti dalle banche dati per le attività di Abstracting and Indexing (A&I). La visione di multipli ed in un certo senso competitivi VAA comporterebbe la partecipazione attiva, ciascuno per il proprio ruolo di competenza, di società professionali, biblioteche nazionali, sull'analisi di dati e metadati. Un esempio di VAA harvesting potrebbe eseguire classificazioni automatiche sui dati o metadati e possibilmente rappresentare percorsi lungo schemi di classificazioni, per esempio nella traduzione di metadati o sistemi di classificazione entro altri linguaggi o per includere il livello educativo nei metadati per la didattica. Un VAA harvesting potrebbe inoltre includere riassunti automatici ed estrazioni di citazioni, annotazioni, ...Un VAA potrebbe essere costruito come parte di un servizio per l'utente finale o potrebbe anche essere un servizio basato sulla macchina per essere usato solo da chi costruisce altri service provider.

Un'area importante è quella che concerne la relazione tra metadati e dati e l'esportazione dei metadati stessi a fini conservativi. Alla fine il gruppo coordinato da Nelson ha discusso sul problema degli oggetti connessi in differenti archivi, in particolare le connessioni a video, datasets. In sostanza un VAA dovrebbe scoprire queste relazioni ed esporre i propri risultati.

L'ultima sessione, coordinata da **Les Grivell**, ha trattato l'editoria scientifica dei periodici non commerciali. Cosa significa esattamente il termine "non commerciale", quali sono i ruoli e le aspettative dei differenti attori nel processo di editoria scientifica: autori, lettori, editori, società accademiche o associazioni e quali i modelli economici più idonei all'incontro delle aspettative di ciascuno dei soggetti coinvolti.

Mirabile l'epilogo finale di **Jean-Claude Guedon** che ha saputo racchiudere i concetti esposti nelle due giornate dai vari relatori in modo intenso e razionale. Nel suo appassionato discorso conclusivo ha sapientemente ricondotto le fila entro le trame complesse che caratterizzano il tessuto organizzativo degli Open Archive, tracciando costellazioni reali con i punti, più o meno luminosi, che segnano il cammino della comunicazione scientifica. Tra questi la crisi dei periodici e degli alti costi di abbonamento e le non chiare responsabilità degli editori commerciali che, negli ultimi cinquant'anni, hanno gestito e trasformato l'editoria scientifica tradizionale entro un'impresa puramente economica orientata esclusivamente al loro profitto. La questione della proprietà intellettuale così centrale e così presente in tutte le discussioni che riguardano la produzione scientifica, ci appare relativamente semplice e chiara, ma la storia ci rivela come le leggi, nel corso del tempo, abbiano creato paradossali estensioni di un diritto naturale, il diritto dell'autore, entro un concetto di proprietà privata che, nel circuito

della comunicazione scientifica, ci appare piuttosto come un ossimoro. In breve, ci dice Guedon, i bibliotecari non devono perdere di vista il fatto che devono aiutare a ricostruire le infrastrutture cognitive corrispondenti alle collezioni aperte dell'era post Gutenberg. Da regista accorto ha saputo offrirci uno scenario che sembra già a portata di mano.