



JLIS.it

Italian Journal of Library and Information Science

Rivista italiana di biblioteconomia, archivistica e scienza dell'informazione

Pubblicazione periodica semestrale (esce in giugno e in dicembre)

ISSN: 2038-5366 (print) – ISSN: 2038-1026 (online)

Website: <http://jlis.it> – Email: [info@jlis.it](mailto:info@jlis.it)

ABBONAMENTO 2014:

*Italia € 50,00; Estero € 60,00.*

SUBSCRIPTION 2014:

*Italy € 50,00; Overseas € 60,00.*

Ledizioni – LediPublishing – Via Alamanni 11

20141 - Milano - Italia

Tel. +39-0245071824 – Fax +39-0242108107 – IVA/VAT: IT04627080965



Vol. 5, n. 1 (Gennaio/January 2014)

# JLIS.it

Italian Journal of Library and Information Science  
Rivista italiana di biblioteconomia,  
archivistica e scienza dell'informazione

Università di Firenze  
Dipartimento di Storia, Archeologia, Geografia, Arte e  
Spettacolo (SAGAS)

**Direttore = Editor in chief**

Mauro Guerrini (Università di Firenze)

**Condirettore = Co-editor**

Gianfranco Crupi (Sapienza Università di Roma)

**Managing editor**

Andrea Marchitelli (CILEA)

**Vicedirettori = Associate editors**

Andrea Capaccioni (Università di Perugia),

Graziano Ruffini (Università di Firenze)

**Staff**

Valentina Demontis, Antonietta Folino,  
Ilaria Fava, Giulia Manzotti, Erika Pasceri

JLIS.it è edita dall'Università di Firenze, Dipartimento di Scienze dell'Antichità, Medioevo e Rinascimento e Linguistica.

PROPOSTE DI PUBBLICAZIONE & PEER-REVIEW: Le submission, compiute tramite il sito web della rivista, verranno inizialmente esaminate da un editor e, superata la prima valutazione, saranno inviate a due revisori per il processo di peer review, al termine del quale verrà notificata l'accettazione o meno del contributo, o l'eventuale richiesta di modifiche.

DIRITTI: JLIS.it applica una licenza "Creative Commons - Attribuzione" (CC-BY) a tutto il materiale pubblicato.

JLIS.it is published by the University of Florence, Department of studies on the Antiquities, Middle Age, the Renaissance and Linguistics.

SUBMISSION & PEER-REVIEW: Papers submitted via the journal website will be checked by one of the editors: if these pass the first step, the papers will be sent to two reviewers for the peer-review process. After this, the author will be notified on the acceptance of his paper, or will be given suggestions on how to improve it.

RIGHTS: JLIS.it is published under a "Creative Commons Attribution License" (CC-BY).





Sommaro  
Table of Contents  
Vol. 4, n. 2 (Luglio/July 2012)

Saggi = Essays

Alberto Salarelli	<i>Right to be forgotten and online newspapers archives: some considerations on social memory in our times</i>	p. 1-20
Giovanna Frigimelica	<i>Business models for selection and acquisition of ebooks for Italian academic libraries. A comparison of three platforms (EBL, ebrary, EBSCO)</i>	p. 21-42
Anna Galluzzi	<i>Libraries and public perceptions: A comparative analysis of the European press. Methodological insights</i>	p. 43-76
Antonella Iacono	<i>From record to data. Linked data and information seeking in the OPAC</i>	p. 77-102
Patrick Lo, Kitty But, Robert Trio	<i>Links between Libraries and Museums: a Case Study of Library-Museum Collaboration at the Hong Kong Maritime Museum</i>	p. 103-120
Maria Teresa Biagetti	<i>Preserving Open Access digital resources</i>	p. 121-148
Gaia Innocenti, Cristina Mugnai, Valdo Pasqui	<i>FLORe – Florence Repository. The University of Florence Open Archive t</i>	p. 149-171
Michele Chieppi, Piera Bergomi, Donatella De Amici	<i>Bibliographic research in nursing science: the sharing experience between Offices at the IRCCS Policlinico S. Matteo di Pavia Foundation</i>	p. 173-208

## Fare il punto = Making the Point on

Maura Funari	<i>Research Data and Humanities: an European overview</i>	p. 209-239
Laura Testoni	<i>MOOCs and academic libraries: a chance or a problem? An overview</i>	p. 241-263
Ben Gu	<i>MARC Formats in China: Local or International?</i>	p. 265-277



# Diritto all'oblio e archivi online dei quotidiani: alcune considerazioni sulla memoria sociale ai nostri tempi

Alberto Salarelli

Un recente provvedimento del Garante della privacy<sup>1</sup>, ci offre l'occasione per affrontare un tema che presenta notevoli profili di interesse sia sul piano teorico (in merito al concetto di memoria sociale) sia su quello pratico (relativamente alle metodologie di archiviazione digitale dei quotidiani). La questione centrale su cui verte il provvedimento concerne il diritto all'oblio. Se esso nasce e si sviluppa nel secolo passato a seguito della necessità di contemperare la libertà di stampa con la tutela della riservatezza e della reputazione, solo in tempi più vicini a noi, e cioè con la diffusione delle reti telematiche, tale diritto assume delle valenze del tutto nuove relativamente alla permanenza, in server accessibili da ciascuno, di dati obsoleti della più diversa natura e tuttavia ancora riferibili a soggetti in vita. Proprio a seguito della facilità di accesso a questi dati il diritto all'oblio si viene sempre più configurando come un risvolto particolare del diritto alla privacy che –come noto– oggi non può più essere inteso, stante l'ottocentesca definizione di Warren e Brandeis, come il semplice diritto a non vedere turbata la riservatezza della propria sfera

<sup>1</sup><http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/2286820>.



personale (*right to be left alone*) ma che invece è da considerarsi come il diritto a mantenere il controllo in merito ai dati che a questa sfera sono da riferirsi, fino al punto da poterne esigere la cancellazione qualora si siano esaurite le finalità per la quale sono stati raccolti e diffusi.

In relazione alla complessità degli ambiti di riferimento, il diritto all'oblio è stato definito come una categoria giuridica "a fattispecie plurima", e questo perché esso

«non riguarda più solo le notizie diffuse dai media e ridiffuse anche quando è venuto meno l'interesse pubblico a conoscerle, ma si estende, più in generale, a coprire ogni informazione o dato che ci riguarda, compresi quelli da noi stessi immessi in rete per le più diverse finalità.»(Pizzetti p. 42)

In questo articolo vogliamo soffermarci su una specifica applicazione di questo diritto, un'applicazione che, per le sue caratteristiche, può essere considerata come una delle fattispecie ontologicamente più "genuine" tanto in relazione alle origini storiche di un così particolare e controverso istituto giuridico (la valutazione in merito alla rilevanza per l'opinione pubblica delle notizie apparse sulla carta stampata è, infatti, all'origine del diritto all'oblio), quanto in riferimento ad un ambito applicativo facilmente circoscrivibile rispetto alla molteplicità dei casi che il mondo digitale può presentare. Ci riferiamo alla possibilità, da parte di un cittadino, di vedere riconosciuto il diritto alla tutela dell'onorabilità della propria persona senza che essa possa venire offuscata dalle ombre di un passato remoto riemerse dalle pagine ingiallite di un quotidiano, pagine improvvisamente riportate a nuova vita grazie al fatto di essere state digitalizzate e rese disponibili online. In altre parole, quali cautele si dovrebbero osservare nell'allestimento di un archivio digitalizzato di un quotidiano per non ledere il diritto di ciascuno a lasciarsi alle

spalle un passato, talora ingombrante, con il quale si sono definitivamente chiusi i conti con la giustizia e, magari, con la propria coscienza?

## 1 Il fatto e il provvedimento

Iniziamo con l'esposizione del fatto all'origine del provvedimento del Garante. Al suo ufficio, nell'ottobre 2012, si è rivolta una persona presentando ricorso nei confronti del gruppo editoriale L'Espresso, e questo perché nell'archivio storico del quotidiano "La Repubblica", consultabile anche attraverso i più comuni motori di ricerca, risultavano presenti tre articoli contenenti dati personali relativi a una vicenda giudiziaria nella quale il ricorrente era stato coinvolto nel passato e, successivamente, prosciolto con formula piena. Nel ricorso si richiedeva, in prima istanza, la rimozione degli articoli e, in subordine, l'aggiornamento delle notizie riportate con la segnalazione della non sussistenza di alcun addebito penale nei confronti dell'interessato, oltre all'adozione di quelle misure tecnologiche necessarie al fine di impedire ai motori di ricerca l'indicizzazione degli articoli in questione.

Il Garante ha accolto il ricorso e, nel suo provvedimento, ha ordinato al gruppo editoriale L'Espresso di predisporre, nell'ambito dell'archivio storico de "La Repubblica", «un sistema idoneo a segnalare (ad esempio, a margine dei singoli articoli o in nota agli stessi) l'esistenza di sviluppi delle notizie relative al ricorrente»; inoltre ha dichiarato il non luogo a procedere in merito alle richieste di interdizione alla consultabilità degli articoli tramite i motori di ricerca esterni al sito Web dell'editore, dal momento che l'editore stesso aveva già provveduto ad adottare le misure tecniche per interdire la loro indicizzazione. Le motivazioni addotte sono, in sostanza, riconducibili a due punti:

1. il trattamento dei dati personali, effettuato al tempo della pubblicazione degli articoli per finalità giornalistiche, all'atto della messa a disposizione in rete dell'archivio digitale del quotidiano risulta compatibile con il trattamento previsto al fine di concretizzare e favorire la libera manifestazione del pensiero e la ricerca storica. Di conseguenza la tutela del diritto allo studio e all'informazione non solo rende insostenibile la richiesta dell'interessato in merito alla rimozione degli articoli dall'archivio storico, ma rende la consultabilità dei medesimi non vincolata dal consenso espresso dell'interessato (cfr. art. 136 e segg. del Codice);
2. d'altro canto l'interessato ha diritto a ottenere l'aggiornamento/integrazione dei dati personali qualora, nel corso del tempo, siano accaduti fatti in grado di incidere in modo significativo sul suo profilo e sulla sua immagine, secondo quanto previsto dall'art. 7 del Codice. A sostegno di questa posizione il Garante, nel suo provvedimento, cita una sentenza della Suprema Corte di Cassazione (n. 5525/2012) relativa ad un caso analogo occorso in un recente passato. In essa si statuisce che

«a salvaguardia dell'attuale identità sociale del soggetto [occorra] garantire al medesimo la contestualizzazione e l'aggiornamento della notizia già di cronaca che lo riguarda, e cioè il collegamento della notizia ad altre informazioni successivamente pubblicate, concernenti l'evoluzione della vicenda, che possano completare o financo radicalmente mutare il quadro evincentesi dalla notizia originaria, a fortiori se trattasi di fatti oggetto di vicenda giudiziaria, che costituisce anzi emblematico e paradigmatico esempio al riguardo.

Quindi, se la vicenda ha avuto ulteriori sviluppi, essi non possono essere ignorati «giacché altrimenti la notizia, originariamente completa e vera, diviene non aggiornata, risultando quindi parziale e non esatta, e pertanto sostanzialmente non vera».

Da quanto esposto appare evidente come il problema di base che il Garante cerca di affrontare consista in un bilanciamento di diritti: da un lato il diritto di cronaca e di critica, esercitabile anche in prospettiva storica, che giustifica la ricercabilità e la consultabilità della notizia seppur a distanza di tempo rispetto a quando si sono svolti i fatti narrati, e dall'altro il diritto della persona di poter tutelare la propria reputazione. In passato tale bilanciamento era, di fatto, garantito dalla difficoltà di accesso all'archivio del giornale e dalle altrettanto complesse procedure di recupero della singola notizia riferita a un determinato soggetto. Nel mondo analogico raggiungere un luogo ove la collezione cartacea (o microfilmata) del quotidiano era conservata e sfogliarne intere annate–pagina dopo pagina–fino all'individuazione della notizia era l'unico modo per farla emergere dal passato. Un modo talmente scomodo da scoraggiare molte potenziali ricerche. Nel mondo digitale tale equilibrio viene ad essere compromesso, vista la facoltà di ciascuno di poter interrogare comodamente da casa propria, in modalità full-text, una gamma sempre più ampia di archivi storici di periodici messi gratuitamente a disposizione dagli editori non solo attraverso i motori di ricerca dedicati a questa specifica funzione e inseriti nei portali delle diverse testate, ma anche attraverso i motori di ricerca generalisti come Google o Bing.

In questa inusitata prospettiva il nuovo mezzo messo a disposizione del pubblico muta in modo sostanziale la funzione dell'archivio stesso. In altri termini, se è vero che, come ha scritto Jacques Derrida, in generale «il senso archiviabile si lascia anche e in anti-

cipo co-determinare dalla struttura archiviante»(Derrida p. 30), nel passaggio dall'analogico al digitale tale "senso", per le caratteristiche di trasparenza e di accessibilità del sistema, si apre con dovizia di risorse sia verso coloro che ambiscono a ricostruire una vicenda attraverso l'applicazione di un'opportuna metodologia storiografica, sia per le indagini nei confronti di cittadini al di sopra (o al di sotto) di ogni sospetto da parte di qualsivoglia individuo animato da intenzioni e curiosità più o meno lecite. Ed allora l'affermazione di Roger Chartier in merito al fatto che «la représentation électronique de l'écrit modifie radicalement la notion de contexte et, du coup, le processus même de la construction du sens» (*Lecteurs et lectures à l'âge de la textualité électronique*) può essere senz'altro applicata anche alle notizie contenute negli archivi online dei quotidiani: in essi si moltiplicano le opportunità di recupero dell'informazione ma, allo stesso tempo, si ampliano i rischi di ricostruzioni storiche false o tendenziose.

Perciò, al di là delle implicazioni concrete del provvedimento (su cui torneremo alla fine di questo scritto), è importante osservare come il Garante colga in pieno la criticità di questa situazione evidenziando le notevoli possibilità offerte dagli archivi digitali dei giornali in merito all'esercizio della libertà riflessiva di informarsi (Gardini pp. 35–38), e, al contempo, gli altrettanto notevoli rischi di intrusioni immotivate nella sfera personale di ciascun cittadino. Ciò che invece non emerge dal provvedimento (ed è giusto che sia così essendo un atto di natura amministrativa) è come il caso in questione rappresenti un esempio concreto e particolarmente significativo di come stia mutando il concetto di memoria sociale nel mondo contemporaneo.

## 2 La memoria sociale e il diritto all'oblio nel mondo digitale

Come il cervello umano sia in grado di trattenere e organizzare i ricordi è questione dibattuta da centinaia di anni. Quello che gli studi più recenti sul suo funzionamento ci hanno dimostrato (Schacter) è che la memoria umana non opera sulla base di schemi classificatori stabili e immutabili: la memoria umana si riscrive continuamente sulla base degli stimoli a cui il soggetto è sottoposto nell'arco della propria esistenza, rafforzando legami reciproci tra i ricordi ma anche eliminando ciò che non risulta più utile alle esigenze del momento. Quindi, al di là di ogni prospettiva meccanicistica, la memoria umana, piuttosto che a un archivio ben ordinato di informazioni, è paragonabile a «un luogo attivo, in permanente ebollizione»(Maldonado p. 120), nel quale le informazioni vengono trattenute, rilasciate o collegate in relazione agli stimoli e alle esigenze del momento. L'oblio, in tale prospettiva, si pone come un processo costruttivo di filtraggio delle informazioni nel senso che «dimenticando, recuperiamo la libertà di generalizzare, concettualizzare e soprattutto agire»(Mayer-Schönberger p. 101); in un certo senso l'oblio, consentendoci di chiudere alcune porte sul nostro passato, ci offre una maggiore libertà di concepire il presente come qualcosa di nuovo da costruire, ci offre—in altre parole—una prospettiva di crescita e di ottimismo verso il futuro, un'aspettativa che manca in quei soggetti afflitti dalla sindrome ipermnestica, incapaci di cancellare i propri ricordi, e perciò inesorabilmente bloccati, nella vita quotidiana, dall'insopportabile zavorra del loro passato, proprio come il personaggio di Funes dell'omonimo racconto di Borges.

Se l'acquisizione su base scientifica di questa modalità operativa della mente umana è cosa recente, noti da tempo sono invece gli ef-

fetti di tale attività, vale a dire che proprio a motivo della fallacia del ricordare l'uomo si è reso conto della necessità di elaborare delle tecniche mnemoniche e, soprattutto, delle forme di «esteriorizzazione della memoria individuale» (Leroi-Gourhan p. 307), utili a superare le proprie manchevolezze fisiologiche. In ultima istanza qualsiasi memoria artificiale ha come obiettivo quella stabilità dei ricordi che, come detto, non è prerogativa della mente: si tratta, insomma, di una forma di compensazione il cui valore è inestimabile proprio in virtù della sua "artificialità" (Ong pp. 119–125), e, pertanto del suo essere iscritta su un supporto materiale esterno rispetto al nostro corpo. Tale esternalità è altresì importante sotto un altro aspetto, e cioè quello relativo all'esperienza collettiva: proprio per il fatto di essere registrata su un oggetto, la memoria esterna è leggibile (e, certo, variamente interpretabile) anche da altri soggetti rispetto a chi l'ha prodotta, configurandosi quindi come elemento strategico per un'esperienza condivisa del ricordare e, quindi, come memoria sociale. In sostanza, come ha scritto Elena Esposito,

«forme di memoria "genuinamente sociali" si sviluppano solo con la disponibilità di *tecnologie della comunicazione*: dalla scrittura e dalla stampa fino ai recenti media elettrici ed elettronici che si rendono via via più indipendenti dalla memoria individuale dei sistemi psichici, e segnano nello stesso tempo degli incrementi nelle prestazioni della memoria». (Esposito p. 19)

Ci perdoni il lettore per queste considerazioni alquanto compendiose relative a problematiche dai vastissimi orizzonti (si è tralasciato, ad esempio, di accennare alle tecniche di memorizzazione dei popoli senza scrittura) ma il nostro scopo è quello di giungere rapidamente all'oggi, e cioè agli effetti della rivoluzione digitale su questi processi.

Viktor Mayer-Schönberger, autore di un importante saggio dedicato

al diritto all'oblio nell'era digitale, sostiene che se nel passato la creazione di una memoria artificiale era una pratica costosa in termini di spesa e di tempo da impiegare per la sua realizzazione—con il risultato che dimenticare è stata la norma e ricordare l'eccezione—con la digitalizzazione, stante i minori oneri di registrazione delle informazioni e la loro accresciuta accessibilità e facilità di recupero, tale situazione si è ribaltata al punto che oggi «il default diventa ricordare e non più scordare»(Mayer-Schönberger p. 59).

La conseguenza di questa nuova prospettiva è che ci troviamo impelagati in una marea di informazioni relative al nostro passato che possono seriamente vincolare il nostro status presente, impedendoci di agire in piena libertà nei confronti di un futuro il cui disegno è pesantemente condizionato dalle tracce che abbiamo lasciato alle nostre spalle. Inoltre la congerie di materiali recuperabile attraverso i più comuni sistemi di information retrieval, leggasi i motori di ricerca, presenta tali livelli di frammentazione e di decontestualizzazione che la tendenza verso un appiattimento della dimensione temporale, rispettosa dell'ordine cronologico di produzione dei documenti, e lo smarrimento del rapporto di autorità tra produttore e contenuto sono più di un rischio paventato, quanto piuttosto una realtà di fatto. Tutto ciò, avverte Mayer-Schönberger, non vuole adombrare i vantaggi delle pratiche di digitalizzazione quanto rilevare che, nel passaggio al mondo nuovo, il senso della memoria subisce una radicale trasformazione della quale è opportuno tenere conto (anche in relazione ai risvolti negativi che essa inizia a manifestare) sia a livello di condotte individuali sia a livello legislativo, con la necessità di emanare opportune normative atte a tutelare il diritto alla riservatezza dei cittadini. Beninteso, l'autore non è così ingenuo da ritenere che l'adozione di un sobrio stile d'uso dei sistemi digitali (fino all'utopica astinenza dai medesimi) e l'inserimento nei codici a tutela della privacy di qualche norma specifica sulla durata della

conservazione dei dati possano essere risolutivi nei confronti di un problema di così vasta portata. La stessa soluzione tecnica da lui proposta—l'inserimento nei file di una data di scadenza—presenta notevoli elementi di opinabilità, non tanto in termini di realizzazione pratica quanto di reale efficacia. Il punto, e il merito principale di Mayer-Schönberger, consiste invece nell'aver posto al centro di un dibattito allargato un diritto, quello all'oblio, che fino ad oggi risultava relegato nelle sfere professionali dei vari esperti della memoria, fossero essi psicologi o sociologi, filosofi o storici, bibliotecari o archivisti. Gli appartenenti a queste ultime due categorie, peraltro, in virtù del loro ruolo di custodi istituzionali della memoria sociale, hanno sollevato da tempo—e cioè fin dagli albori della rivoluzione digitale—un paio di questioni che l'autore austriaco sembra considerare solo marginalmente e che, invece, rivestono un'importanza strategica.

La prima questione riguarda le tecniche di preservazione a lungo termine dei documenti digitali: a differenza di quanto sembra credere Mayer-Schönberger, non è affatto certo che essi siano destinati a durare per un tempo immemorabile. Esempi neanche troppo remoti ci hanno insegnato che i rischi di perdita dei dati conservati su supporti elettronici sono elevatissimi e che, al di là dei bassi costi iniziali di digitalizzazione, «conservare documenti digitali costa molto, ma molto di più, che conservare documenti cartacei»(Zanni Rosiello p. 100).

La seconda questione concerne la possibilità di recuperare in modo efficace le informazioni dalle basi dati ove esse sono potenzialmente contenute. Se è indubbio che gli strumenti di information retrieval si sono evoluti in modo impressionante in termini di rapidità nello scandagliare archivi sempre più smisurati, siamo ancora molto lontani dalla prospettiva di una rete globale nella quale i documenti risultino semanticamente correlati l'uno all'altro. Escludendo al-

cune risorse del Web ove i materiali inseriti risultano controllati e classificati in modo analitico, la massa critica dell'intero sistema è unicamente percorribile attraverso il riconoscimento delle parole chiave inserite come stringa di ricerca. Il che, come noto, può essere sufficiente per il recupero di dati fattuali ma non certamente per una ricerca efficace in termini di indagine attorno ad un determinato soggetto o argomento. La prospettiva di una memoria digitale sempre più estesa, stabile e imperitura deve quindi essere temperata con la presenza di fattori di rischio strutturali (l'invecchiamento dei supporti materiali e l'obsolescenza dei formati di digitalizzazione) e funzionali (la definizione di strumenti efficaci di utilizzo dei dati) che rappresentano altrettante cause di un possibile oblio dell'informazione nell'era digitale. Abbiamo a disposizione archivi sempre più ampi di cui tuttavia non conosciamo la capacità di durare nel tempo e per i quali non siamo certi di possedere le giuste chiavi per poterli aprire e consultare.

C'è, infine, un'ultima questione che è da considerarsi come specifica prerogativa di chi opera professionalmente nell'ambito della mediazione dell'informazione e, in specifico, nel mondo dell'archivistica: ci riferiamo all'atteggiamento, consolidatosi nel corso del tempo, relativamente alle procedure di selezione e scarto dei documenti. Nell'era della documentazione analogica tali attività scaturivano dal compromesso tra le esigenze di conservazione integrale degli archivi e la fattibilità pratica di sopperire a queste occorrenze (Lodolini p. 214). Nel mondo digitale, secondo la prospettiva di Mayer-Schönberger, il paradosso che legava nel passato conservazione e distruzione verrebbe ad esaurirsi, stante la disponibilità immensa di memoria per poter trattenere tutto. In realtà, come si è detto, le cose stanno in altro modo: la mera conservazione del digito binario, di per sé niente affatto scontata, non significa garantire nel futuro la consultabilità dei documenti. Quindi, anche per la contem-

poraneità, la selezione rimane un'attività dolorosa ma strategica per tentare di trasmettere alle generazioni future quei documenti che, con un atto di necessario arbitrio, sono stati ritenuti meritevoli, in virtù della loro significanza, di una preservazione ad altri negata. Come ha scritto Stefano Vitali, «ciò deriva non tanto o soltanto da ragioni di economia di spazio e di costi, quanto dal semplice fatto che non può esistere trasmissione della conoscenza e della memoria, che non implichi, al contempo, variabili dosi di oblio: la memoria, anche quella documentaria, è per sua natura selettiva»(Vitali p. 180). Dal momento che, anche nella dimensione digitale, conservare significa scegliere e distruggere, nessuno nega come su tali pratiche di selezione possano influire interessi non coincidenti con i criteri di scarto definiti dalla comunità professionale. Il fatto che il costituente o il gestore di un archivio elettronico abbiano la possibilità di alterare il sistema secondo le loro più o meno lecite esigenze non fa altro che confermare, anche per la contemporaneità, la validità della teoria foucaultiana sul potere detenuto da queste figure. L'archivio, come noto, deve intendersi non solo come sedimento dell'attività operativa dei soggetti produttori, ma anche come autorappresentazione, cioè come modo di trasmettere un'immagine di sé (Zanni Rosiello p. 36) che, malgrado il vincolo archivistico, può essere la risultante di molteplici attività di edulcorazione nella scelta dei documenti e nel modo di relazionarli reciprocamente. Ciò premesso, quel che conta rilevare è l'esistenza, nell'ambito degli archivi pubblici, di diritti da tutelare che si traducono in regole da rispettare: si segna in tal modo un confine tra il lecito e l'illecito che condiziona in modo determinante la trasmissione della memoria, sia nel modo di organizzare i documenti e la loro accessibilità, ma anche nei confronti di ciò che la collettività, tramite la sovranità esercitata attraverso la funzione legislativa, ritenga debba essere conservato o distrutto. In merito al caso di specie che stiamo affrontando, è altresì interes-

sante osservare come la normativa sulla privacy (cfr. in particolare gli artt. 101 e segg. del Codice) estenda il proprio ambito di applicazione anche agli archivi privati: il potere del Garante di intervenire sulle forme costitutive dell'archivio digitale di un quotidiano, impedendo la rimozione di articoli e obbligando il costituente a gestire le informazioni in modo specifico, è un esempio significativo di come la memoria digitale non sempre sia il frutto di scelte (o non scelte) casuali o di parte. In altri termini, gli archivi pubblici, o aperti al pubblico, i luoghi istituzionali della preservazione della memoria sociale, sono altra cosa dai sedimenti documentari sparsi nelle vaste praterie della rete ove non vige il diritto codificato ma, come nel selvaggio West, la legge della pistola più veloce o più potente. Gli archivi, come tutti i beni culturali, sono una testimonianza di civiltà, ed è per questo motivo che il trattamento dei documenti in essi contenuti non può che riflettere i legittimi interessi della collettività, e non solo quelli del singolo cittadino, a salvaguardia della propria vita privata: «l'autonomia individuale non può spingersi fino a pretendere la cancellazione della storia, una ripulitura del nostro passato da tutto ciò che percepiamo come sgradevole, fondando così un diritto di autorappresentazione che obblighi tutti gli altri a vederci solo come noi vogliamo»(Rodotà p. 65).

### **3 Gli archivi online dei quotidiani e la tutela della privacy: fra contestualizzazione e accessibilità**

Queste ultime considerazioni attorno agli archivi controllati dalle istituzioni delegate dalla collettività alla preservazione della memoria sociale ci riportano al caso da cui siamo partiti. Infatti il provvedimento del Garante, con il suo obbligo per il gestore dell'archivio di

predisporre un idoneo sistema di aggiornamento/integrazione delle notizie in esso contenute, tocca una questione teoretica di portata non indifferente.

È indubbio che uno degli elementi che distinguono l'organizzazione degli archivi analogici rispetto a quelli digitali sia il diverso grado di contestualizzazione che distingue i primi dai secondi. L'esperienza comune insegna come i risultati della ricerca di documenti effettuata in rete porti sovente a risultati non correlati tra loro sulla base del vincolo archivistico, ma unicamente in relazione alla presenza in essi della medesima stringa di ricerca digitata dall'utente. In queste aggregazioni documentarie *on the fly*, come ha osservato Federico Valacchi, «vengono meno i rapporti di univocità tra soggetto produttore e sedimentazione documentaria»(Valacchi p. 36) con il rischio di suggerire interpretazioni fuorvianti dei fatti accaduti, delle responsabilità ad essi correlati e della reputazione dei soggetti coinvolti. Questo spiega, da un lato, il divieto imposto dal Garante nei confronti di un'accessibilità indiscriminata degli archivi dei quotidiani da parte dei motori di ricerca e, dall'altro, l'obbligo di individuare un sistema di contestualizzazione interno agli archivi stessi che dia ragione dei successivi sviluppi della vicenda. Il problema sorge nel momento in cui tale enunciazione, di massima in merito alla predisposizione di un idoneo sistema di aggiornamento/integrazione, debba essere tradotta in pratica. Lasciando da parte le non irrilevanti questioni relative all'effettiva realizzabilità concreta, sul piano informatico, di un sistema di questo tipo, ci si chiede con quali criteri il gestore dell'archivio dovrebbe provvedere all'attualizzazione delle notizie in esso contenute. E, soprattutto, se sia lecito che l'attività di ricostruzione di un fatto–fino ad ora prerogativa dello storico–debba essere istituzionalmente praticata da chi invece gestisce i documenti. In altre parole, chi sarebbe titolato a perseguire *ex post* la ricerca di quella "verità" della notizia a cui fa riferimento la sentenza n.

5525/2012? Il produttore dell'archivio o il ricercatore? Pare lecito obiettare che «Le eventuali integrazioni potrebbero dunque divenire oggetto di sindacato circa la correttezza e pertinenza dei richiami e della descrizione postuma degli eventi. Il preteso aggiornamento e la contestualizzazione finirebbero così con sostituire alla semplice certezza e chiarezza della raccolta di archivio una potenzialmente controversa interpretazione ricostruttiva» (Mantelero p. 160).

Insomma: nel suo comprensibile desiderio di tutelare il diritto alla privacy del singolo cittadino il Garante, sulla scorta della discussa sentenza della Cassazione (la quale—detto per inciso—si spinge ben oltre imponendo, addirittura, la cancellazione della notizia qualora i fatti presenti non siano tali da “assurgere a rilevanza storica”), sembra avere imboccato una strada quantomeno sdruciolevole dal punto di vista della teoria archivistica e storiografica, oltre che discutibile in merito all'efficacia pratica dell'obiettivo atteso. Non è in discussione, lo ripetiamo, la legittimità di un intervento normativo sulle forme di costituzione degli archivi, quanto piuttosto i limiti a cui tale intervento dovrebbe attenersi. Certo, un sistema di contestualizzazione non verrebbe a ledere il vincolo archivistico rappresentato, in un giornale, dalla sequenza cronologica di pubblicazione degli articoli e dalla loro organizzazione in numeri e annate, ma potrebbe sminuire la sua importanza, suggerendo salti verso alcune notizie correlate a discapito di altre possibili e altrettanto importanti e contestualizzanti connessioni. Si dirà: questo è il corrispettivo necessario per mantenere bilanciati i diritti del singolo e della collettività nella nuova dimensione digitale. Tuttavia non ci pare che l'attribuzione all'archivista di compiti che, tradizionalmente, sono prerogativa dello storico possa rappresentare un passo avanti efficace in questa direzione.

L'aspetto su cui, a nostro avviso, si poteva intervenire con maggiore efficacia è un altro, vale a dire la regolamentazione dell'accesso al-

l'archivio. Infatti la più agevole consultabilità dell'archivio digitale rispetto a quello cartaceo è l'elemento sostanziale che fa scattare la necessità di bilanciare i differenti diritti: se nessun giudice si è mai sognato di ordinare la messa in opera di idonei sistemi di aggiornamento/integrazione delle notizie contenute nelle collezioni cartacee dei quotidiani, e tantomeno di "ritagliare" da essi quelle notizie non meritevoli di «assurgere a rilevanza storica», ciò è dovuto, lo ripetiamo, alla difficoltà di accesso a tali documenti. Quindi la strada più semplice, anche per gli archivi dei quotidiani online, consiste nel predisporre degli sbarramenti tali da rendere la loro consultabilità meno agevole di quanto non sia oggi: su questa linea, del resto, si colloca l'obbligo imposto dal Garante di interdire l'indicizzazione degli articoli da parte dei motori di ricerca generalisti. Ma si potrebbe fare di più, ad esempio prevedendo qualche forma di identificazione dell'utente remoto, al quale si richiederebbe di leggere e rispettare le norme contenute nel codice di deontologia e di buona condotta per il trattamento di dati personali per scopi storici. Nessuno, certo, si illude che ciò possa rappresentare una soluzione definitiva al problema: la storia di Wikileaks ci insegna come in rete le fughe di notizie (e la loro conseguente replica su mille altri siti) siano un fenomeno difficilissimo da controllare, a maggior ragione se l'archivio non venga violato informaticamente ma i documenti siano fatti circolare da qualche talpa. Tuttavia la compilazione di un modulo online e l'invio di una copia del proprio documento d'identità sarebbero passaggi più che sufficienti a interdire le ricerche di molti perditempo e, in ogni caso, ad allertare tutti coloro che consultano l'archivio in merito alle conseguenze civili e penali della divulgazione di notizie remote per le quali oggi non siano più validi quei profili di interesse pubblico che ne avevano giustificato la pubblicazione nel passato. Insomma: se è vero che nel mondo degli archivi cartacei «l'esistenza di vincoli all'accesso

può costituire una condizione per garantire il bisogno di riservatezza dei vari portatori di interessi e impedire che tale necessità sia soddisfatta attraverso la manipolazione, la distruzione o la dissimulazione» (Giuva pp. 166-167) non c'è motivo per cui tali vincoli non possano sussistere anche in ambiente digitale. Vincoli blandi, certo: in fondo il caso di cui ci stiamo occupando non riguarda dossier secretati ma notizie già pubblicate sulla carta stampata. Vincoli che, a nostro modo di vedere, potrebbero rappresentare un primo ma non trascurabile passo verso una didattica all'uso dell'archivio online della testata. Se infatti, oltre alla registrazione, fossero messi a disposizione degli utenti dei canali di contatto con i responsabili dell'archivio, molti dubbi sulla liceità e sulle condizioni d'uso delle informazioni recuperabili dalle pagine del passato potrebbero essere facilmente dissipati. Ma stiamo parlando in ipotesi visto che due delle tre grandi testate italiane con un archivio online ("La Repubblica", "Corriere della Sera" e "La Stampa") non forniscono alcun contatto con i responsabili degli archivi stessi. Solo il quotidiano torinese indica un indirizzo di posta elettronica. Forse, nell'era dei social network, si potrebbe fare uno sforzo in più. Gli archivisti italiani si sono resi conto che da qualche anno le sale di studio sono frequentate da un pubblico di tipo nuovo, composto da persone con interessi e pratiche di ricerca estremamente eterogenei e caratterizzato da un rapporto con le fonti sensibilmente diverso rispetto agli storici di professione (Vitali p. 92): instaurare un dialogo si è rivelato il primo, imprescindibile passo per poter svolgere un servizio efficace di mediazione tra documenti e utenti. Crediamo che anche verso gli utenti degli archivi online, un popolo dal profilo ancora più variegato, tale riconoscimento reciproco si renda necessario.

## Riferimenti bibliografici

- Chartier, Roger. *Lecteurs et lectures à l'âge de la textualité électronique*. 2001. [http://cv.uoc.edu/~04\\_999\\_01\\_u07/chartier2.html](http://cv.uoc.edu/~04_999_01_u07/chartier2.html). (Cit. a p. 6).
- Derrida, Jacques. *Mal d'archivio. Un'impressione freudiana*. Napoli: Filema, 2005. (Cit. a p. 6).
- Espósito, Elena. *La memoria sociale. Mezzi per comunicare e modi di dimenticare*. Roma-Bari: Laterza, 2001. (Cit. a p. 8).
- Gardini, Gianluca. *Le regole dell'informazione. Principi giuridici, strumenti, casi*. Milano: Bruno Mondadori, 2005. (Cit. a p. 6).
- Giuva, Linda. «Archivi e diritti dei cittadini». *Il potere degli archivi. Usi del passato e difesa dei diritti nella società contemporanea*. A cura di Linda Giuva, Stefano Vitali e Isabella Zanni Rosiello. Milano: Bruno Mondadori, 2007. 135–201. (Cit. a p. 17).
- Leroi-Gourhan, André. *Il gesto e la parola*. Torino: Einaudi, 1965. (Cit. a p. 8).
- Lodolini, Elio. *Archivistica. Principi e problemi*. Milano: Franco Angeli, 1990. (Cit. a p. 11).
- Maldonado, Tomás. *Memoria e conoscenza. Sulle sorti del sapere nella prospettiva digitale*. Milano: Feltrinelli, 2005. (Cit. a p. 7).
- Mantelero, Alessandro. «Il diritto all'oblio dalla carta stampata ad Internet». *Il caso del diritto all'oblio*. A cura di Franco Pizzetti. Torino: Giappichelli, 2013. 145–172. (Cit. a p. 15).
- Mayer-Schönberger, Viktor. *Il diritto all'oblio nell'era digitale*. Milano: Egea, 2013. (Cit. alle pp. 7, 9).
- Ong, Walter J. *Oralità e scrittura. Le tecnologie della parola*. Bologna: Il Mulino, 1986. (Cit. a p. 8).
- Per la protezione dei dati personali, Garante. *Archivi storici on line dei quotidiani e reperibilità dei dati dell'interessato mediante motori di ricerca esterni*. 2013. Registro dei provvedimenti n. 31 del 24 gennaio 2013.
- Pizzetti, Franco. «Il prisma del diritto all'oblio». *Il caso del diritto all'oblio*. A cura di Franco Pizzetti. Torino: Giappichelli, 2013. 21–63. (Cit. a p. 2).
- Rodotà, Stefano. *La vita e le regole. Tra diritto e non diritto*. Milano: Feltrinelli, 2006. (Cit. a p. 13).
- Schacter, Daniel L. *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Boston: Houghton Mifflin, 2001. (Cit. a p. 7).
- Valacchi, Federico. *La memoria integrata nell'era digitale. Continuità archivistica e innovazione tecnologica*. Corazzano: Titivillus, 2006. (Cit. a p. 14).
- Vitali, Stefano. *Passato digitale. Le fonti dello storico nell'era del computer*. Milano: Bruno Mondadori, 2004. (Cit. alle pp. 12, 17).

Zanni Rosiello, Isabella. *Gli archivi nella società contemporanea*. Bologna: Il Mulino, 2009. (Cit. alle pp. 10, 12).

ALBERTO SALARELLI, Università degli Studi di Parma.

[alberto.salarelli@unipr.it](mailto:alberto.salarelli@unipr.it)

<http://salarelli.altervista.org/>

---

Salarelli, A. "Diritto all'oblio e archivi online dei quotidiani: alcune considerazioni sulla memoria sociale ai nostri tempi". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #14. DOI: [10.4403/jlis.it-14](https://doi.org/10.4403/jlis.it-14). Web.

**ABSTRACT:** Inspired by a recent decision of the Italian Data Protection Authority, the article examines the issues related to the right to be forgotten in relation to the newspapers archives available on the Internet. This particular case is presented as a concrete example of the need to find a form of balance between the right to privacy and the right to historical research; it is also a particularly significant example of how the concept of memory is changing in the contemporary world. The article ends with a proposal for a more controlled access to online newspapers archives, not for the purpose of censoring the documents contained in them, but to inform users about the legal consequences of their activity.

**KEYWORDS:** Right to be forgotten; Online newspapers archives; Right to privacy; Freedom of research; Social memory.

---

Submission: 2013-09-10

Accettazione: 2013-10-05

Pubblicazione: 2014-01-01





# Modelli di acquisto di e-book per le biblioteche dell'università italiana. Tre piattaforme a confronto (EBL, Ebrary, Ebsco)

Giovanna Frigimelica

## 1 Introduzione

Negli ultimi anni la diffusione degli e-book ha avuto una crescita esponenziale. L'accesso 24 ore su 24, la portabilità, gli strumenti avanzati offerti dalle piattaforme rendono questo prodotto particolarmente interessante in ambito accademico. L'attenzione è alta anche da parte dei bibliotecari per i vantaggi organizzativi (potenziale aumento dello spazio fisico a disposizione nelle biblioteche, migliore gestione della collezione, riduzione dei costi di *stack management*). Le biblioteche delle università si trovano ad attraversare una fase di sperimentazione per quanto concerne la selezione, l'acquisizione e gli altri processi collegati alla fruizione degli e-book da parte degli utenti (per un inquadramento generale si veda Morris e Sibert). Queste attività, al momento, risultano essere molto complesse a causa di alcuni fattori, primo fra tutti la frammentazione del mercato e la mancanza di standard a livello di licenze. Attualmente



esistono diverse piattaforme per l'accesso perpetuo<sup>1</sup> o il prestito di e-book, di singoli editori o di aggregatori, caratterizzate da diversi business model, funzionalità più o meno avanzate, ricchezza dei titoli presenti in catalogo. A ciò si aggiungono fattori collegati alle scelte editoriali, tra cui il ritardo, che va dai 3 ai 18 mesi (Waltersp. 191), a seconda della previsione di uscita di versioni *hardcover* e *paperback*, sul rilascio della versione elettronica, le limitazioni sull'uso dell'e-book (numero di accessi contemporanei, possibilità di stampa, copia/incolla etc.) e infine l'effettiva disponibilità dei titoli.

Questo articolo affronta le modalità di accesso agli e-book offerte da tre piattaforme (EBL, Ebrary, EBSCO) in particolare con l'ottica di stabilire se un'università statale italiana di grandi dimensioni, che offre un ampio panorama di corsi di laurea (di primo livello e specialistici/magistrali), master di primo e secondo livello, dottorati di ricerca e altre attività didattiche e formative, possa acquistare titoli stranieri in versione elettronica anziché cartacea, e a che condizioni. Viene analizzato il catalogo offerto da queste tre piattaforme, i business model che vi soggiacciono, le funzionalità disponibili. Il taglio è pratico-sperimentale, in quanto i paragoni tra le piattaforme si basano su un set di 150 monografie, pubblicate negli ultimi 5 anni da editori diversi e acquistate in cartaceo, utilizzato come campione rappresentativo. Non vengono affrontati altri aspetti sicuramente importanti, come la catalogazione, il prestito e la promozione presso gli utenti, l'impatto sulle abitudini di lettura o di studio da parte di studenti, docenti e ricercatori, le problematiche collegate alla gestione dei diritti digitali (DRM).

---

<sup>1</sup>Non è del tutto corretto, anche se molto diffuso, l'uso del termine "acquisto" in quanto nella maggior parte dei casi il file fisicamente resta nella piattaforma del fornitore, e la biblioteca ha diritto ad accedervi secondo modalità di volta in volta definite.

## 2 Le piattaforme: criteri per una scelta

Nel momento in cui si intraprende una valutazione tra piattaforme per individuare quella più confacente, si deve tenere conto di alcuni importanti fattori (Grigson; Roncevic; Vasileiou, Hartley e Rowley), tra cui si segnalano:

- tipologia (per editore, aggregatore, distributore etc.);
- specificità di pubblico a cui si rivolge (tutte le biblioteche, in prevalenza pubbliche, in prevalenza accademiche e di ricerca etc.);
- numero di titoli presenti, numero e tipologia di editori rappresentati;
- frequenza di aggiornamento e di crescita della collezione di titoli disponibili;
- lingue e argomenti che la collezione copre o privilegia;
- business model per l'acquisto e/o il prestito (possibilità di acquisto perpetuo, accesso per singolo utente, per numero di utenti in contemporanea, accesso illimitato, prestito breve etc.);
- costi (sia dei titoli, in funzione delle modalità di accesso/possesso, sia -quando applicati- della piattaforma);
- funzionalità tecniche (opzioni di ricerca, compatibilità con il gestionale di biblioteca in uso e/o l'OPAC-discovery tool, possibilità di stampare, copiare, annotare etc.);
- formati disponibili dei file (ePub, PDF, HTML etc.);

- funzionalità gestionali avanzate (statistiche COUNTER, disponibilità record MARC etc.).

Partendo da una prima schematizzazione delle piattaforme esistenti che si rivolgono principalmente al mercato bibliotecario universitario, esse possono essere suddivise in grandi gruppi:

- di editori: usualmente offrono solo i titoli pubblicati dall'editore in modalità sia di accesso perpetuo sia di abbonamento a collezioni
- di distributori: trattano sia libri cartacei sia elettronici di un numero ristretto di editori
- di aggregatori: contrattano con un ampio numero di editori e vendono gli e-book secondo modelli di vario tipo.

Gli aggregatori garantiscono la più ampia copertura di titoli, editori e argomenti, nonché diverse modalità di accesso ai contenuti. La scelta è stata di concentrarsi quindi su tre piattaforme di aggregatori affermate in ambito universitario, individuate tra quelle più utilizzate nel mondo anglosassone («Ebook usage in U.S. academic libraries»; Waltersp. 189),<sup>2</sup> distribuite o rappresentate in Italia, ovvero Ebook Library (EBL)<sup>3</sup>, Ebrary<sup>4</sup> ed Ebsco<sup>5</sup>.

---

<sup>2</sup>Anche Overdrive rientra in questa categoria ed è molto diffusa, ma da contatti intercorsi con il distributore italiano è emersa una più spiccata attenzione al mercato delle biblioteche pubbliche e di istruzione superiore, l'impossibilità inoltre di avere accesso in prova alle funzioni di amministratore (compresi i prezzi dei titoli a seconda delle licenze) ha portato ad escludere questa piattaforma dall'analisi.

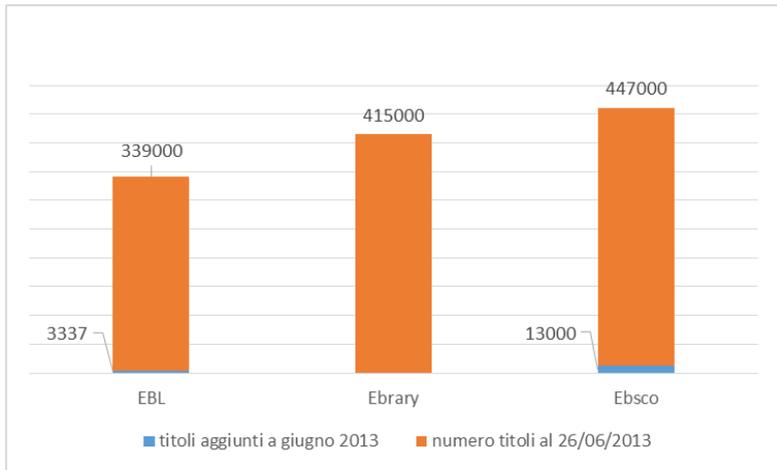
<sup>3</sup><http://www.ebllib.com>.

<sup>4</sup><http://www.ebrary.com/corp>.

<sup>5</sup><http://www.ebscohost.com/ebooks/about>.

### 3 I contenuti

Come prima cosa è stata analizzata l'offerta in termini di catalogo, reputando essenziale verificare il numero di titoli disponibili nella piattaforma e il numero di editori rappresentati.<sup>6</sup>

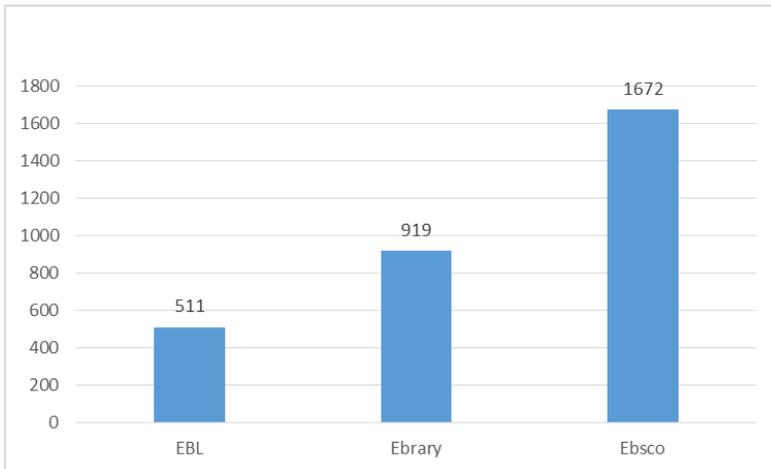


**Figura 1:** Numero titoli presenti per piattaforma

Per estrapolare il numero di titoli, riepilogati in figura 1, è stata lanciata una ricerca "a vuoto" ed è stato riportato il numero di risultati ottenuti (arrotondati al migliaio). Nel caso di Ebsco sono stati applicati i filtri "ebooks+ Downloadable Content" per eliminare altre tipologie di materiale diverse dalle monografie; allo stesso modo per Ebrary la ricerca è stata filtrata per tipologia materiale "books". Sono stati estrapolati anche i titoli aggiunti nel corso del

<sup>6</sup>Tutte le ricerche sono state svolte tra maggio e giugno 2013. I cataloghi di queste banche dati sono in continuo aggiornamento, per elaborare i totali complessivi si è scelta come data unica per tutte e tre le piattaforme il 26 giugno.

mezzo di giugno per valutare il tasso di crescita della collezione (nel caso di Ebrary non è stato possibile recuperare il dato, EBL ed Ebsco consentono di affinare la ricerca per data di aggiunta del titolo a catalogo). Il numero di editori presenti per EBL è desunto dal sito del distributore italiano (Cenfor) che ne riporta l'elenco<sup>7</sup>, per Ebrary è stato dichiarato, mentre per Ebsco è stato desunto dalla piattaforma<sup>8</sup> che consente il browsing per editore.



**Figura 2:** Numero editori presenti per piattaforma

Le differenze, come emerge dai grafici riportati, sono significative, in particolare, per gli editori. Ebsco, infatti, propone nel proprio catalogo oltre il triplo degli editori di EBL. Il numero di titoli, invece, non si discosta con la stessa proporzione (Ebsco ha il 24% di titoli in più di EBL), pertanto è ragionevole dedurre che parte degli editori presenti in Ebsco abbia un numero contenuto di titoli a

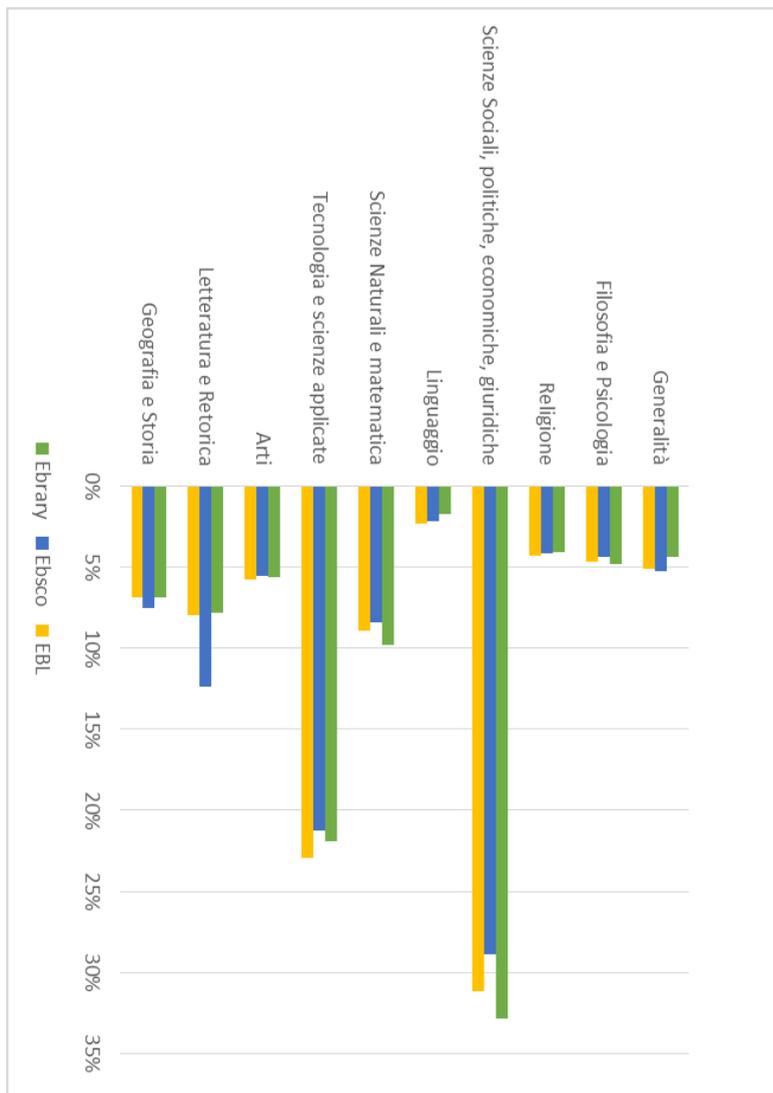
<sup>7</sup><http://www.cenfor.net/download/EBLPublishers.pdf>.

<sup>8</sup>Accesso al catalogo tramite EBSCOhost Collection Manager (ECM).

catalogo. L'informazione sul numero di editori qui fornita è quantitativa e non entra nel merito qualitativo, in quanto "l'importanza" può variare a seconda della mission o della specializzazione di una biblioteca rispetto a un'altra. I principali grandi editori scientifici sono mediamente rappresentati nelle tre piattaforme.

Le piattaforme analizzate si rivolgono principalmente a biblioteche delle università e di enti di ricerca, pertanto gli e-book disponibili saranno in massima parte testi di ricerca e di supporto alla didattica. Dando per assunto il target di riferimento, è utile valutare quale copertura venga assicurata per argomento dai titoli a catalogo. È stata condotta una ricerca applicando il filtro per classe Dewey. Nel caso di EBL ed Ebsco la somma dei titoli per classe è leggermente più alta del totale dei titoli a catalogo in quanto è possibile che ad alcuni testi sia stata attribuita più di una classe. Viceversa nel caso di Ebrary la Classificazione decimale Dewey è stata assegnata solo a 232 mila titoli (circa la metà di quelli disponibili), pertanto il dato fornito va considerato una proiezione. Come emerge dal grafico in figura 3 nella pagina successiva, tutte le piattaforme hanno una copertura maggiore per le classi 300 e 600.

Altro dato di interesse, in particolare per un cliente italiano quale può essere la biblioteca di un ateneo del nostro paese, è la copertura linguistica. Dalle ricerche compiute nei cataloghi delle tre piattaforme emerge che il 90% dei titoli proposti è in inglese. In minima parte sono presenti e-book in tedesco, mentre sono pressoché assenti quelli in francese, spagnolo e *–last but not least–* italiano. Anche in questo caso è stata lanciata una ricerca a vuoto, selezionando come tipologia di materiale solo gli e-book, applicando quindi il filtro per lingua della risorsa e riportando il numero di risultati ottenuti, espressi in percentuale nel grafico rappresentato in figura 4 a pagina 29.



**Figura 3:** Titoli per classe Dewey

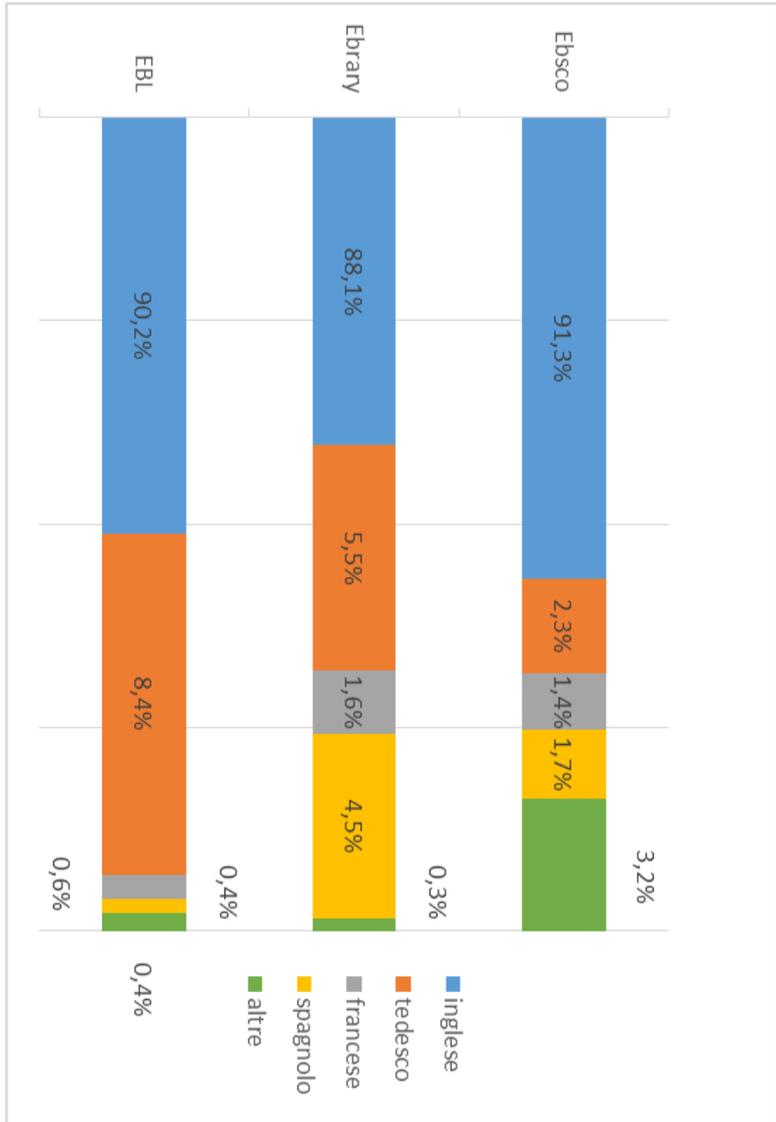


Figura 4: Titoli per lingua

## 4 Business model

Dopo aver valutato il contenuto presente nelle piattaforme, il passo successivo è valutare i business model offerti (Besen e Nataraj Kirby; Sharp e Thompson, «E-books in academic libraries: lessons learned and new challenges»; «'Just in case' vs. 'Just in time': e-book purchasing models»; Vasileiou, Hartley e Rowley. 23-26). Essi si possono dividere in quattro grandi gruppi:

1. acquisto di titoli singoli (con un numero limitato o illimitato di accessi simultanei);
2. abbonamento a pacchetti o *subject collections*;
3. *short-term loan*, cioè noleggio dell'e-book per un periodo di tempo limitato (normalmente espresso in giorni, con accesso in modalità *one/single user*) al termine del quale il titolo non sarà più accessibile salvo acquisto con altra modalità da parte della biblioteca;
4. *patron (o demand) driven acquisition (PDA)* che consente, dietro attivazione e settaggio di alcuni parametri da parte della biblioteca (ad es. il budget massimo a disposizione), l'acquisto in automatico di e-book che non fanno parte della collezione della biblioteca ma sono presenti nel catalogo, e che abbiano avuto un certo utilizzo da parte dell'utenza (come lettura di almeno 10 pagine, consultazione per almeno 10 minuti etc.).

Nella tabella 1 nella pagina successiva sono schematizzati i business model disponibili per piattaforma.

	<b>EBL</b>	<b>Ebrary</b>	<b>Ebsco</b>
<b>offerta di pacchetti/collezioni</b>	no	si, in abbonamento (accessi illimitati)	si, in abbonamento (accessi illimitati)
<b>possibilità noleggio a breve termine (one/single user)</b>	si (1-7-14-28 gg)	si (1-7 gg, max 3 prestiti per titolo)	si (1-7-14-28 gg)
<b>patron-driven acquisition</b>	si	si	si

**Tabella 1:** business model disponibili per piattaforma

Mentre Ebrary<sup>9</sup> ed Ebsco<sup>10</sup> consentono l'accesso agli e-book con modalità sostanzialmente simili (*one/single user, 3 users, unlimited multi-users*), EBL<sup>11</sup> propone un sistema parzialmente diverso, chiamato *non-linear lending*. Ogni e-book viene venduto con una "dotazione" di 325 giorni/prestiti nel corso dell'anno (12 mesi a partire dall'acquisto). I prestiti possono avvenire anche in simultanea, il numero di giorni di ciascun prestito viene sottratto dal numero totale di giorni autorizzati per il libro. Esauriti i 325, bisognerà attendere la scadezza dei 12 mesi per avere diritto ad altri (o acquistare una seconda copia del libro). Esistono inoltre altre due opzioni, oltre al *short-term loan*:

- *textbook*: e-book considerati come libro di testo o materiale didattico da parte degli editori, vengono forniti con i 325 giorni/prestiti e un accesso limitato a un massimo di 3 utenti contemporanei;
- *unlimited access*: disponibile soprattutto per i titoli di reference, consente un numero illimitato di accessi per un numero illimitato di giorni all'anno.

I prezzi per l'acquisto di titoli singoli dovrebbero essere fissati dall'editore<sup>12</sup> per il *one/single user* e quindi uguali in tutte le piattaforme.

---

<sup>9</sup>Alcune informazioni sui business model sono disponibili a <http://www.ebrary.com/corp/models.jsp>.

<sup>10</sup>Alcune informazioni sui business model sono disponibili a <http://www.ebscohost.com/ebooks/acquisition-options>.

<sup>11</sup>Alcune informazioni sui business model sono disponibili a [http://www.eblib.com/docs/Ebook%20Library%20QA\\_2007.pdf](http://www.eblib.com/docs/Ebook%20Library%20QA_2007.pdf).

<sup>12</sup>Ebrary dichiara: «Pricing per title is at publisher list price for single-user access, and 150% of list price for unlimited multi-user access where available»(<http://www.ebrary.com/corp/models.jsp>). EBL dichiara: «Libraries pay list price for ebook content and a platform fee or annual maintenance fee for the EBL platform. Ebook prices are set by the publishers, and we recommend that they set these at same-as-print levels»([http://www.eblib.com/docs/Ebook%20Library%20QA\\_2007.pdf](http://www.eblib.com/docs/Ebook%20Library%20QA_2007.pdf)). EBSCO

A seconda del numero di utenti simultanei per titolo i prezzi variano da piattaforma a piattaforma. Il costo del *short-term loan* dipende dalla durata selezionata e dal prezzo dell'e-book nella piattaforma, e si attesta tra il 10 e il 30%. Questa opzione può essere utile, in particolare, in sostituzione del prestito interbibliotecario. Normalmente è disponibile solo in modalità *one/single user*. Collezioni e pacchetti offerti in abbonamento hanno il vantaggio di avere un costo complessivo minore rispetto all'acquisto titolo per titolo, nonché delle modalità di accesso più ampie (spesso multiutente). Lo svantaggio è l'assenza di controllo sulla politica delle collezioni (nel pacchetto saranno inclusi inevitabilmente anche titoli non di interesse) e l'eventuale perdita di accesso ai contenuti, qualora non fosse possibile negli anni successivi sostenere il costo dell'abbonamento.

EBL	Ebrary	Ebsco
3.000 \$ in anticipo, o +10% sul costo dell'e-book fino ad arrivare a un massimo di 5.000 \$; 750\$ manutenzione annuale	hosting fee annuale per possesso perpetuo in caso di acquisto di e-book, +5% sul costo dell'e-book, con un minimo di 250 \$ e un massimo di 1.500 \$	assenti

**Tabella 2:** Costi fissi (IVA esclusa) per la piattaforma

## 5 Funzionalità

Gli aspetti tecnici e le funzionalità delle piattaforme sono importanti per valutare fino in fondo che servizio potremo offrire ai nostri

dichiara: «With no markups and no fees on any title, EBSCO eBooks is a cost-effective option for any library» (<http://www.ebscohost.com/ebooks/acquisition-options>).

utenti una volta selezionati gli e-book. La maggior parte delle piattaforme è accessibile con i principali browser e consente la lettura anche on line sul computer. Bisogna considerare, infatti, che non tutti gli utenti avranno un *portable reading device* personale su cui leggere l'e-book. Tutte e tre le piattaforme analizzate utilizzano Adobe Digital Editions per leggere offline il testo completo. Alcune funzionalità servono a rendere l'esperienza di lettura più efficace per l'utente (la possibilità di sottolineare e commentare il testo, copiarne e stamparne delle parti, l'integrazione con dizionari e la possibilità di ricercare parole all'interno del testo etc.). Altre funzionalità sono utili dal punto di vista gestionale-amministrativo per i bibliotecari. La presenza di statistiche serve a valutare, ad es. in termini di costo per uso, il reale impatto della collezione di testi elettronici sull'utenza, anche in paragone con le statistiche di circolazione dei volumi cartacei (Sharp e Thompson, «'Just in case' vs. 'Just in time': e-book purchasing models»). L'integrazione con l'OPAC-discovery tool in uso aiuta a promuovere il servizio e a consentire una ricerca semplificata dei testi di interesse. Nella tabella 3 a fronte sono schematizzati i principali aspetti tecnici per piattaforma.

## 6 Analisi della disponibilità e dei costi di un set di titoli

Per valutare in maniera più approfondita il possibile utilizzo di queste piattaforme per l'acquisto di e-book in ambito accademico italiano, è stato individuato un set di 150 monografie<sup>13</sup> acquistate

---

<sup>13</sup>In altri studi simili (in particolare Link; Pomerantz) il set di titoli era più alto (poco più di 500) ma l'indagine era svolta nell'ambito di biblioteche di università americane, dove la lingua di riferimento è l'inglese. Nel caso di un'università italiana è evidente che una buona parte delle acquisizioni consti di traduzioni o titoli in lingua italiana. Pertanto si ritiene che il campione di titoli, pur essendo più basso, sia

	<b>EBL</b>	<b>Ebrary</b>	<b>Ebsco</b>
<b>ricerca full-text</b>	Si	Si	si
<b>copia/incolla</b>	si (max 5% del testo)	max 60 stringhe copiate per libro per sessione	max 60 pagine per libro per sessione
<b>stampa</b>	si (max 20% del testo)	max 60 pagine per libro per sessione	max 60 pagine per libro per sessione
<b>durata max del prestito di e-book acquistati dalla biblioteca</b>	periodi multipli di 7 (min 1 max 28 gg) definiti dalla biblioteca	in base alla licenza di accesso (max 7 gg per il single user e il 3 users, 14 per il multi)	definita dal bibliotecario (in ore, giorni, settimane)
<b>rest. prima della scadenza</b>	No	Si	no
<b>prenotazione tit. in prestito</b>	si	No	si
<b>formato file</b>	PDF, ePub	PDF, ePub	PDF, ePub
<b>record MARC</b>	si	si	si
<b>Open URL/Z39.50</b>	si	Si	si
<b>COUNTER compliant</b>	si	Si	si
<b>autenticazione federata</b>	Athens, Shibboleth	Athens, Shibboleth	Athens, Shibboleth

Tabella 3: Aspetti tecnici e principali funzionalità delle piattaforme

in cartaceo dalle biblioteche dell'Università di Cagliari negli ultimi 6-8 mesi, con le caratteristiche sotto indicate volte a garantire la rappresentatività del set:

- testi prevalentemente in inglese, con una piccola percentuale di titoli campione in tedesco e francese (la scelta è stata dettata dalla copertura linguistica dei cataloghi offerti nelle piattaforme);
- titoli appartenenti ad aree disciplinari diverse e varie (biomedico scientifico, tecnologico, scienze umane, scienze sociali-giuridico-economiche);
- data di pubblicazione negli ultimi 5 anni (con qualche titolo campione di anni precedenti);
- testi pubblicati da vari editori (evitando di concentrare quindi la ricerca sui principali grandi editori scientifici).

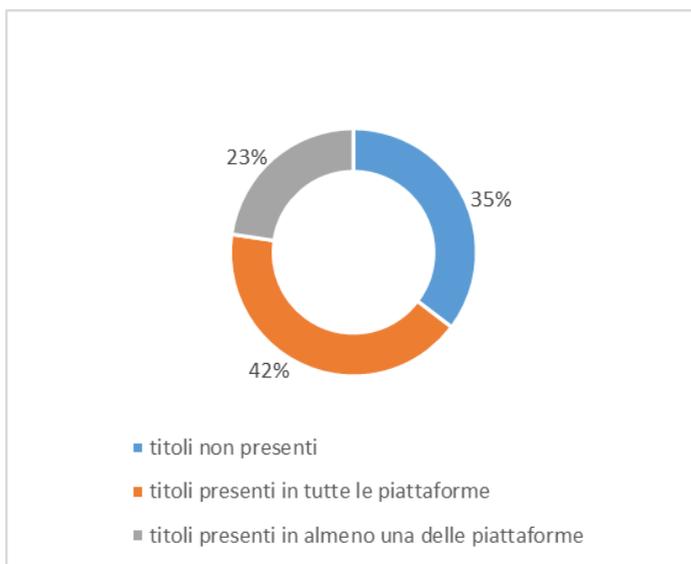
I titoli così individuati sono stati ricercati nelle tre piattaforme: 53 non sono stati trovati in nessuna delle tre piattaforme analizzate<sup>14</sup>, 63 sono presenti in tutte e i restanti 34 sono presenti in almeno una (v. figura 5 nella pagina successiva). Ebrary si è dimostrata la piattaforma con più titoli del set (dettaglio in figura 6 a fronte).

Per poter effettuare un paragone in merito ai costi, si è scelto di comparare il prezzo per l'acquisto dell'e-book titolo per titolo, tralasciando modalità quali il *short-term loan* o il *pay-per-view* nonché l'abbonamento a collezioni. Per EBL è stato preso come riferimento il prezzo per il *non-linear lending* (a volte disponibile solo nella modalità *textbook* e *unlimited access*). Per EBL ed Ebsco sono stati

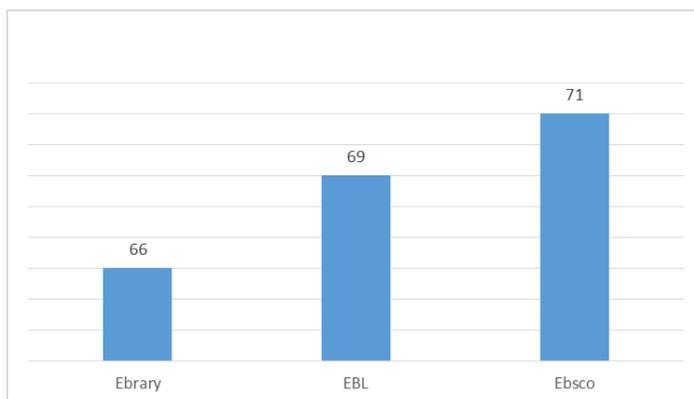
---

ugualmente significativo per la realtà italiana.

<sup>14</sup>Non è stato verificato se l'editore fornisce informazioni sull'esistenza di una versione elettronica del testo, quindi i titoli mancanti potrebbero essere disponibili solo in versione cartacea.



**Figura 5:** Titoli rispetto al set



**Figura 6:** Titoli non presenti per piattaforma rispetto al set

selezionati i prezzi per il *one/single user*. In questo confronto non si è tenuto conto di eventuali costi aggiuntivi di accesso o manutenzione relativi alla piattaforma.

Per 39 titoli (pari al 62%) sui 63 del set presenti in tutte le piattaforme c'è coincidenza di prezzo.<sup>15</sup> Dato che i tre aggregatori, come indicato sopra, dichiarano che il prezzo è fissato dall'editore, si aspettava una percentuale pari al 100%.

Per valutare quanto l'Ateneo avrebbe speso acquistando la versione elettronica al posto della cartacea, è stato confrontato il prezzo pagato per il volume cartaceo (espresso in euro e ribassato del 20% rispetto al prezzo di copertina<sup>16</sup>) e il prezzo degli e-book (espresso in dollari nelle piattaforme, pertanto è stato convertito in euro<sup>17</sup> e maggiorato del 21% di IVA, tassa applicata al momento in Italia su questo prodotto). Non sono state aggiunte eventuali altre maggiorazioni anche se è noto che alcune commissionarie, per il mercato librario cartaceo, calcolano uno spread a titolo di parziale rimborso delle spese e commissioni bancarie italiane ed estere.<sup>18</sup> Mediamente il costo dell'e-book è più alto rispetto al cartaceo del 124% per Ebrary, del 128% per Ebsco e del 153% per EBL. Differenze più marcate si riscontrano nel caso in cui sia stato possibile acquistare l'edizione paperback del testo anziché l'hardcover, arrivando fino ad un aumento del prezzo per l'e-book del 500%. Nel caso specifico del nostro set, questa possibilità si ravvisa più frequentemente nei titoli

---

<sup>15</sup>In 11 casi in cui il titolo è presente in Ebsco ed Ebrary con prezzo coincidente e in EBL con prezzo più alto, quest'ultima piattaforma fornisce il titolo come textbook o unlimited, consentendo pertanto un numero di utenti/accessi più ampio delle altre due piattaforme.

<sup>16</sup>Le monografie vengono acquistate per tramite di una commissionaria che si è aggiudicata la gara assicurando la fornitura con lo sconto del 20%.

<sup>17</sup>Al cambio di riferimento del 26/06/2013 rilevato dalla Banca d'Italia, ovvero 0,7678, si veda a [http://www.bancaditalia.it/banca\\_centrale/cambi/rif/2013/06/cambi\\_rif\\_2606.htm](http://www.bancaditalia.it/banca_centrale/cambi/rif/2013/06/cambi_rif_2606.htm).

<sup>18</sup>Ad esempio la DEA migliora del 3%, <http://www.deastore.com/help/260>.

del settore scienze umane e scienze sociali-economico-giuridiche, meno nei settori biomedico-scientifico e tecnologico. Probabilmente ciò si spiega con la necessità in questi ultimi settori di acquistare con tempistiche più strette le novità editoriali senza attendere la disponibilità dell'edizione paperback.

## 7 Conclusioni

Appare evidente, sia dalla letteratura citata sia da quanto emerge dall'analisi svolta, che non è possibile sostituire completamente l'acquisto di titoli stranieri cartacei con la corrispondente versione elettronica. I titoli in lingua italiana, inoltre, sono in numero estremamente esiguo, fattore che condiziona inevitabilmente l'utilizzo che queste piattaforme possono avere nel processo di selezione e acquisizione dei testi da parte di una università italiana. Andando oltre il set specifico di titoli individuati per questa ricerca e le piattaforme analizzate, secondo altri studi effettuati in area anglosassone solo la metà dei titoli acquistati da biblioteche accademiche è disponibile in e-book (Walters p. 190-191). L'embargo, inoltre, a cui le corrispondenti versioni elettroniche sono attualmente sottoposte (3-18 mesi) limita ulteriormente l'abbandono completo del cartaceo, in particolare per quelle discipline in cui è strettamente necessario accedere tempestivamente al contenuto o all'edizione aggiornata.

La diversa copertura di titoli, editori e argomenti, nonché le diverse licenze di accesso e utilizzo, sono fattori che rendono difficile scegliere un canale di acquisto unico, motivo per cui la maggior parte delle biblioteche che sta sviluppando politiche di gestione di collezioni di e-book si affida a più fornitori e a diversi modelli, a seconda della disponibilità e del titolo specifico. La presenza di costi fissi per la piattaforma può limitare ulteriormente la scelta, oltre a far aumentare il costo finale per l'acquisto di questi prodotti.

L'abbonamento a collezioni o a pacchetti predeterminati dal fornitore risponde alla necessità di un contenimento dei costi garantendo la massa critica di e-book da offrire agli utenti, ma ha come svantaggi la presenza di titoli non pertinenti e la possibilità che, in assenza di fondi costanti, si perda l'accesso ai contenuti.

È necessario che editori e distributori trovino un business model che consenta loro di mantenere i profitti ma venga maggiormente incontro ai bisogni delle biblioteche delle università e dei loro utenti. Allo stato attuale i bibliotecari devono necessariamente entrare attentamente nel merito delle condizioni che queste piattaforme pongono per l'utilizzo, dei costi sul lungo periodo, dei benefici in relazione alle statistiche d'uso, valutando caso per caso, con un importante investimento in termini non solo economici ma anche di risorse umane e di politica delle collezioni.

## Riferimenti bibliografici

- Besen, Stanley M. e Sheila Nataraj Kirby. *E-books and libraries: An economic perspective*. Chicago: American Library Association, 2012. (Cit. a p. 30).
- «Ebook usage in U.S. academic libraries». *Library journal*. (2012). <<http://www.thedigitalshift.com/research/ebook-usage-reports/academic>>. (Cit. a p. 24).
- Grigson, Anna. «An introduction to e-book business models and suppliers». *E-books in libraries: a practical guide*. A cura di Kate Price e Virginia Havergal. London: Facet, 2011. 19–36. <<http://digirep.rhul.ac.uk/items/bf558385-a28c-b465-582e-0e9da35be076/1>>. (Cit. a p. 23).
- Link, Forrest E. «Are we there yet? An analysis of e-book equivalent coverage in highly-circulated titles at The College of New Jersey Library». *Collection Building* 31. (4 2012): 132–135. (Cit. a p. 34).
- Morris, Carolyn e Lisa Sibert. «Acquiring e-books». *No shelf required : e-books in libraries*. A cura di Sue Polanka. Chicago: American Library Association, 2011. <[http://www.alaeditions.org/%20No\\_Shelf\\_Chapter\\_6](http://www.alaeditions.org/%20No_Shelf_Chapter_6)>. (Cit. a p. 21).
- Pomerantz, Sarah. «The availability of e-books: examples of nursing and business». *Collection Building* 29. (1 2010): 11–14. (Cit. a p. 34).
- Roncevic, Mirela. «E-book Platforms for Libraries». *Library Technology Reports* 49. (3 2013). (Cit. a p. 23).
- Sharp, Steve e Sarah Thompson. «E-books in academic libraries: lessons learned and new challenges». *Serials* 22. (2 2009): 136–140. (Cit. a p. 30).
- . «'Just in case' vs. 'Just in time': e-book purchasing models». *Serials* 23. (3 2010): 201–206. (Cit. alle pp. 30, 34).
- Vasileiou, Magdalini, Richard Hartley e Jennifer Rowley. «Choosing e-books: a perspective from academic libraries». *Online Information Review* 36. (1 2012): 21–39. (Cit. alle pp. 23, 30).
- Walters, William H. «E-books in Academic Libraries: Challenges for Acquisition and Collection Management». *portal: Libraries and the Academy* 13. (2 2013): 187–211. <[http://www.press.jhu.edu/journals/portal\\_libraries\\_and\\_the\\_academy/portal\\_pre\\_print/current/articles/13.2.walters.pdf](http://www.press.jhu.edu/journals/portal_libraries_and_the_academy/portal_pre_print/current/articles/13.2.walters.pdf)>. (Cit. alle pp. 22, 24, 39).

GIOVANNA FRIGIMELICA, Università degli Studi di Cagliari.

[g.frigimelica@unica.it](mailto:g.frigimelica@unica.it)

<http://www.frigimelica.it>

---

Frigimelica, G. "Modelli di acquisto di e-book per le biblioteche dell'università italiana. Tre piattaforme a confronto (EBL, Ebrary, Ebsco)". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #8986. DOI: [10.4403/jlis.it-8986](https://doi.org/10.4403/jlis.it-8986). Web.

**ABSTRACT:** In recent years, the spread of ebooks has grown exponentially. 24/7 availability, portability, the advanced tools offered by platforms make this product particularly interesting for academic libraries, which are in an experimental phase with regard to the selection and acquisition of ebooks. The article analyzes three platforms (EBL, Ebrary, EBSCO), to find out whether an Italian medium-sized university library can buy foreign titles in electronic rather than paper, and under what conditions. It provides information about titles catalog, business model, features available. The cut is practical/experimental, since the comparison among the platforms is based on a set, used as a representative sample, of 150 monographs, published in the last 5 years by different publishers and purchased on paper.

**KEYWORDS:** E-books, Library collections

**ACKNOWLEDGMENT:** L'autore desidera ringraziare Andrea Marchitelli, Nicola Cavalli e Laura Testoni per il confronto e l'amichevole lettura del testo, Licosa, Ebsco publishing e Cenfor per la disponibilità, e infine i colleghi dell'Università di Cagliari che hanno collaborato fornendo informazioni e dando consigli.

---

Submission: 2013-07-03

Accettazione: 2013-07-28

Pubblicazione: 2014-01-01





# Libraries and public perceptions: A comparative analysis of the European press. Methodological insights

Anna Galluzzi

## 1 Premise

Measuring and assessing the quality of library services is an integral part of the managerial approach to libraries (Solimine). As a consequence, where a patron-driven approach is more rooted and library is conceived as a social institution (Traniello) and managed as a complex organisation, the need to adopt assessing tools in order to explicit the results gained by library services was felt earlier and more intensely. In recent times, the mainly quantitative approach has turned to a properly qualitative one and tools like interviews and focus groups have been adopted to let patrons' expectations and perceptions about libraries and their services coming out (Pickard; Faggiolani). As a consequence of using the methods of the social sciences (Bryman; Corbetta) and moving from the relative objectivity of quantitative data to the substantial subjectivity of opinions, the level of complexity of evaluation has grown, as well as the capacity to understand the phenomena which affect the library life. Under this

methodological framework, library quality in all its meanings and sides was first emphasised; later the analysis has gone beyond the boundaries of the library environment. Faced with the deep changes occurred in the social, economic and technological context, libraries need to prove their usefulness in increasing the social wealth and be sure this is recognised by the society as a whole. Over the last years there has been much research and discussion about the future of libraries (Coffman; Bell; Lankes R.; “America’s Libraries for the 21st Century: An annotated bibliography”; Osif, “W(h)ither libraries? The future of libraries, part 1”; “W(h)ither libraries? The future of libraries, part 2”), particularly public libraries (Di Domenico; Galluzzi, “Che ne sarà dell’impero bibliotecario?”; “Biblioteche pubbliche tra crisi del welfare e beni comuni della conoscenza. Rischi e opportunità”; *Biblioteche per la città. Nuove prospettive di un servizio pubblico*; Levien; Salarelli; Conti); the numerous papers and contributions range from the analysis of new models and new perspectives of library services in the digital era to the investigation of the cultural, political and sociological premises libraries are based upon. The economic crisis has exacerbated this situation as libraries are facing increasing budget cuts and forced closures.<sup>1</sup> Consequently, the issue of the future of libraries has gone beyond the boundaries of the librarians’ playground, flooding the ground of public opinion by means of the mass media. In these difficult times, many methods have been applied to prove the social and economic impact of libraries (Poll, “I dati che abbiamo e i dati che potremmo avere: nuovi sviluppi nella valutazione delle biblioteche”; *Bibliography “Impact and outcome of libraries”*; Chiessi; Ventura, *La biblioteca rende: Impatto sociale e economico di un servizio culturale*; “La valutazione

---

<sup>1</sup>The following are just a couple of examples of websites born to debate the budget cuts and the library closures in the UK, to publish news about these issues and to advocate for libraries: <http://www.publiclibrariesnews.com> and <http://www.voicesforthelibrary.org.uk>

della biblioteca pubblica: problematiche e strumenti di misurazione dell'impatto culturale, sociale e economico"; Markless and Streetfield) and, in many cases, the resulting reports have contributed to putting libraries on the political agenda. Nonetheless, the methods used so far are mainly library-centred and not far enough from the library world to catch on the prevailing and indefinite feelings permeating society. This is why the analysis of newspapers has been figured out as an alternative method to measure the relevance and the public perception of libraries (Thelwall; Giménez Guardiola and Pedreño). Looking at newspapers is an attempt to move outside the specialised literature concerning libraries. Approaching newspapers for this type of research requires awareness of the strengths and weaknesses of this method which are a consequence of the specific characteristics newspapers have as sources of information and of their relation with the public opinion building process (Barsamian and Chomsky; Habermas). Consequently, it is essential to outline how newspapers work (Zelizer; Boldrini; Lorusso and Violi; Martin and Copeland; Sheridan Burns). Newspapers make their news selection starting mainly from the national and international press agencies. This means that the decision to publish a piece of news depends on the characteristics and nature of the newspaper and is by itself an aspect to be analysed and interpreted. The selection of news published on newspapers depends on the issues which are considered of paramount interest for their readership or those to be put to the attention of the public opinion. Newspapers are crucial in defining the political agenda as they are important loudspeakers of the main opinion leaders, groups, lobbies and movements and deeply affect the public perception. On the other hand, it is undeniable that nowadays traditional newspapers are not the one and only conveyors of public opinions. Their initial role in informing citizens and spreading information has been complemented and partially

replaced by other communication media: first radio and television, then the Internet. From the points of view of promptness and audience, the traditional newspapers have been completely walloped by the other media. Nonetheless, they have reacted reinforcing their role in commenting, debating and analysing the news, which has been further boosted by the launch of the 2.0 newspapers' websites. In sum, though newspapers are not completely representative of the public opinion anymore, they are an important detector of the issues at stake and their non specialised - even sometimes inaccurate - approach is particularly useful in bringing out stereotypes and general opinions. Consequently, they can offer an interesting point of view, particularly on smaller subjects like libraries.

## **2 Aims and contents of the research**

This research aims at quantifying and qualifying the presence of issues concerning libraries in the European press over the last years (Faggiolani; Bryman; Pickard) in order to answer the following research questions:

- which are the most discussed topics concerning libraries and have they changed over the last years?
- are there any significant differences between the European countries in the debate about libraries?
- are there any significant differences between the European newspapers in the debate about libraries?

The chronological span covered by the research is five years, from 2008 to 2012. This choice was made because 2008 is generally considered the starting point of the economic crisis which is still deeply affecting the Western economies and political scenarios. As for the

geographical coverage, the focus has been put on Europe, the so-called "old continent", which is currently observed with special attention by the Western countries and the other nations based on capitalistic economies, as it is the main theatre where the economic crisis is deploying its heavy consequences. The countries taken into account are the United Kingdom, France, Spain and Italy, since they are considered representative of different areas and cultural traditions in Europe, as well as expression of diverse economic and political situations. A second selection was made among the numerous print newspapers published, with the objective of choosing two titles for each country according to the following basic criteria. The two newspapers were picked among those of national relevance, the most widespread and the oldest in each country, avoiding - if possible - those officially representing political parties and the radical ones. Attention was also put in covering both the centre-left area and the centre-right, in order to have different points of view on similar issues and more elements to explain the weight each newspaper gives specific topics in relation to governmental policies. The economic and sport newspapers were not taken into account. The selected newspapers are the following:

- The United Kingdom: *The Times* (centre-right) and *The Guardian* (centre-left)
- France: *Le Figaro* (centre-right) and *Le Monde* (centre)
- Spain: *El Mundo* (centre-right) and *El País* (centre-left)
- Italy: *Corriere della Sera* (centre/liberalism) and *La Repubblica* (centre-left)

As for Italy, while *La Repubblica* evidently belongs to the centre-left area, *Il Corriere della Sera* cannot be considered a right-wing newspaper; it rather is an expression of liberalism and, from a political

point of view, is closer to the centre area. France has a similar situation, since *Le Monde* can be classified as belonging to the centre-area, whereas the left is represented by more radical titles. The databases used to go through the chosen newspapers were Lexis Nexis™ News for *Le Monde* only and Factiva™ for all the other titles. For all the newspapers - whenever possible - the source selected was the national print edition (without the local editions). The coverage guaranteed by the same database for different newspapers is not always comparable, depending upon the agreements with the publishers. In sum, all the titles cover the national print edition and encompass the supplements; two of them, *El País* and *Corriere della Sera*, include the regional or local editions too. In one case, *La Repubblica*, though the database declares that regional editions are available as separate sources, the fall in the number of articles retrieved from 2011 on (see figure 1 on page 50) makes clear that, for the previous period (2008-2010) the database encompassed the pages of the local editions as well. This lack of consistency cannot be corrected, as it is not possible to clearly distinguish the articles coming from the national edition and those included in the local ones. This is why the comparisons between the absolute numbers of articles for each country and newspaper were avoided and a statistical analysis of the data was needed to make the numbers meaningful. The keywords used as query parameters in the full text search were "librar\*" and "bibliot\*" (according to the language of the newspaper). For the French newspapers only, a second query was performed using "mediat\*" as keyword in the full text search, in order to avoid losing the articles concerning public libraries (which are mainly called "médiathèques" in France).

### 3 The numbers

The articles retrieved using the above-mentioned parameters are 41,611. After the retrieval of the articles responding to the query parameters, the second step was to select the pertinent ones, i.e. those articles which concern libraries in a proper sense. The uses of the terms "library" and "librarians" which were not considered pertinent are the following:

- personal libraries;
- house furniture;
- the place where an event or an exhibition is held;
- the place where a murder or a crime occurs;
- the title of a book series;
- a bunch of physical or digital objects (e.g. iTunes as a music digital library);
- the name of sections of newspapers;
- idiomatic expressions and figures of speech encompassing the term "library", like "Library of Babel" or "bookworm" (which is "library mouse" in French, Italian and Spanish);
- libraries cited in tender announcements;
- libraries cited in biographies and stories.

The pertinent articles are 3,659.

	2008		2009		2010		2011		2012		Overall total	
	Total	Pertinent	Total	Pertinent								
Newspapers	915	78	1070	103	1306	84	1428	104	1992	268	6711	637
Corriere della Sera	3584	231	3445	228	3104	215	529	43	627	92	11289	809
La repubblica	4499	309	4515	331	4410	299	1957	147	2619	360	18000	1446
Italy	1162	110	1161	68	1009	75	961	82	769	99	5062	434
The times	1395	63	1349	77	1251	125	1269	131	851	110	6115	506
The guardian	2557	173	2510	145	2260	200	2230	213	1620	209	11177	940
The United Kingdom	630	62	517	55	446	48	436	27	478	37	2507	229
Le figaro	433	62	487	86	455	91	408	59	448	44	2231	342
Le monde	1063	124	1004	141	901	139	844	86	926	81	4738	571
France	1065	81	1220	107	1306	110	1348	110	1015	126	5954	534
El Pais	384	40	382	42	367	30	308	32	301	24	1742	168
El Mundo	1449	121	1602	149	1673	140	1656	142	1316	150	7696	702
Spain	9568	727	9631	766	9244	778	6687	588	6481	800	41611	3659
Overall total												

Figure 1: Overall numbers of the articles examined and pertinent ones

## 4 Processing the articles

After the selection, a text and content analysis of the articles was carried out (Neuendorf and Copeland; Krippendorff). To support this process many computer software for qualitative data analysis are available, which help in organizing and analyzing non-numerical and unstructured data. As for this research, the pros and cons of using this type of software were attentively evaluated. Though aware of the many advantages (speed, completeness, objectivity and precision) of an automatic processing, the risk to think that the whole analysis could be delegated to computer software, instead of using them to speed up and enhance it, was given a special credit. In fact, without a previous coding process, the automatic search of keywords throughout the texts was deemed to be potentially misleading. Without the appropriate questioning during the analysis, the risk either to run the wrong procedures and queries or to create fictitious relationships among concepts and documents was considered high. Furthermore, in this case the analysis of articles in four different languages would have increased the variables and the potential inconsistency of the results. All things considered, the analysis was carried out manually and no text analysis software was used, starting from the firm belief that no software can replace human reasoning. A certain degree of subjectivity was considered somewhat inevitable and acceptable, provided that the reasoning is logical and intelligible and the results are consistent. First of all, each article was identified with a univocal name and an Excel<sup>TM</sup> worksheet was prepared to host the results of the coding. Then, the articles were analysed and coded. Coding written texts presents nearly the same problems as subject indexing and classification, though in a less formalised context (Neuendorf and Copeland; Krippendorff). In particular, the following main aspects of the text and content analysis were taken into consideration:

- the different ways of conveying the same concepts (synonyms and quasi-synonyms);
- the need to distinguish among different types of concepts (main concepts, their attributes and so on);
- the need to clarify the relationships among concepts (for example, parent/children).

This is why Text analysis is a gradual process, as a clearer map of concepts and their relationships can be approached only by means of subsequent attempts. At the beginning, the texts were carefully reviewed and all concepts and ideas were annotated as they appeared and then grouped. Afterwards, the articles were coded according to the following categories and sub-categories:

- country
- newspaper title
- year of publication
- prevalence or not of libraries as subject of the article
- type of library considered: Public, National, Academic, School, Special/Specialised, No specification or more than one type
- main topic of the article: Mission/Roles, Conservation/Holdings/Catalogue, Digitisation/Digital libraries, History, Reading/Marketing, Politics/Strategy/Management, Library closures/Budget cuts, Internet/E-book/Technology, Services/Users, Staff/Recruitment, New libraries/New buildings, Acquisitions/Open access, Buildings/Architecture.

- the newspaper section where the article is published: Opinions/Letters/Debates, Culture/Education, In brief, Cities/Local news, National news, World/International news, Market/Economy/Business, Society, Science, Other
- in case the article deal with libraries of other countries, in addition to the other codes, it is coded under the country it talks about, according to the following scheme: Portugal, Spain, Austria/Switzerland, Scandinavian Countries, Eastern Europe, Greece, Germany, Italy, France, UK/Ireland, Europe in general, Rest of the World/World in general.

As for the topics, which are the more thorny aspect of the coding, they were defined insofar as the text analysis went on; in the end, a flat schema (not a classification schema) was adopted, putting together those issues which could be assimilated and trying to avoid any overlapping. While some groupings were obvious and look self-evident (e.g. "Mission/Roles", "Politics/Strategy/Management"), others, like "Conservation/Holdings/Catalogue" or "Acquisitions/Open Access", need a further explanation as they came from an articulated reasoning. As for the former, it should be stressed that newspapers rarely deal with library catalogues in a technical way, rather they consider catalogues as a direct expression of a rich library collection.

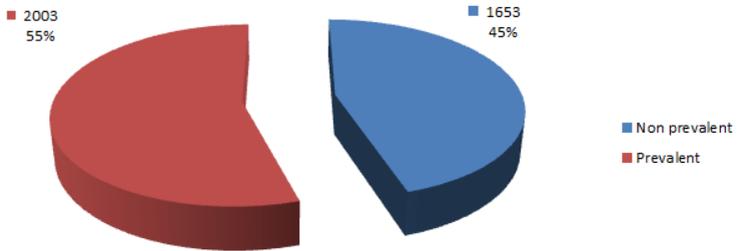
Therefore, the "catalogue" issue is just another way to celebrate what libraries hold and preserve. As far as the "Acquisitions/Open Access" topic is concerned, the label was chosen to put together all those issues pertaining access to library collections. The attention of newspapers is not about selection methods and tools (which would be a too technical approach), but about variety and accessibility of the collection which encompass books and periodicals acquisitions and alternative ways of making them accessible.

To correctly interpret the results of the subject analysis, events and debates which have come out and raised prolonged attention during the considered period should be highlighted, as they could have inspired a high number of articles and created a full strand, which could be either supra-national or strictly linked to the national context. Another clarification is needed for the geographical groups used to categorise the articles concerning libraries of other countries. Like the whole research, they follow a Europe-centred approach. Consequently, all the articles talking about countries outside Europe are put together under the label "Rest of the World/World in general", whereas the European countries are processed individually or are grouped according to geographical and cultural areas. After processing all the articles and completing the Excel<sup>TM</sup> worksheet, various pivot tables were built to sort and summarise the data and these were the starting point for the discussion of the findings.

## 5 Some first findings

The data analysis and processing is ongoing and the research findings are far from being complete. As for now, some first findings in form of tables and graphs can be offered with some first comments. A preliminary quantitative analysis of the articles retrieved concerns the prevalence or not of the library subject (fig. 2 on the next page). Actually, many articles only cite libraries within descriptions and debates regarding more general issues and together with many other topics. These articles are anyhow important, as they vouch for the consideration of libraries when talking about issues like cultural policies, budget planning, national heritage and so on. However, the articles totally centred on libraries are much more relevant in respect to this research, since they highlight the issues that the newspapers deem to be of general interest and valuable of becoming subject of

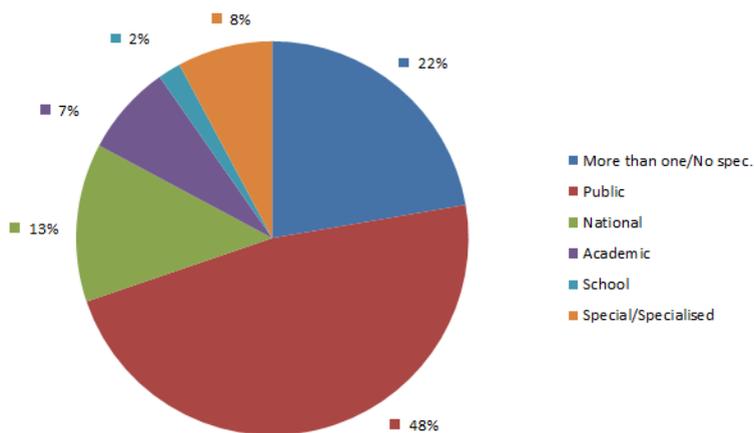
the public debate. In addition, these articles allow the researcher a deeper examination of topics and offer more material for the textual analysis.



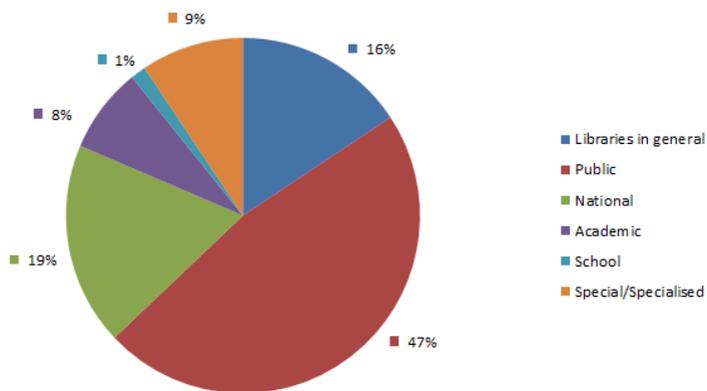
**Figure 2:** Articles where the library subject is prevalent vs. those where it is not

Another question this research has tried to answer is which types of libraries newspaper articles deal with the most. In figures 3 on the following page and 4 on the next page the results of the analysis are shown respectively as percentages of the total number of articles and as percentages of those articles which mainly talk about libraries.

According to figure 3 on the following page, almost half of the total articles are devoted to public libraries. The second most represented type of libraries is national libraries (13%), followed by special and specialised libraries (8%) and academic libraries (7%). School libraries are quite absent from the chosen newspapers. In the 22% of the articles the subject are libraries in general or more types of library altogether. Considering only the articles where library subject is prevalent (fig. 4 on the next page), the percentage of articles talking about libraries in general is lower, whereas the percentage of



**Figure 3:** Types of libraries (% on total articles)



**Figure 4:** Types of libraries (% on articles where the library subject prevails)

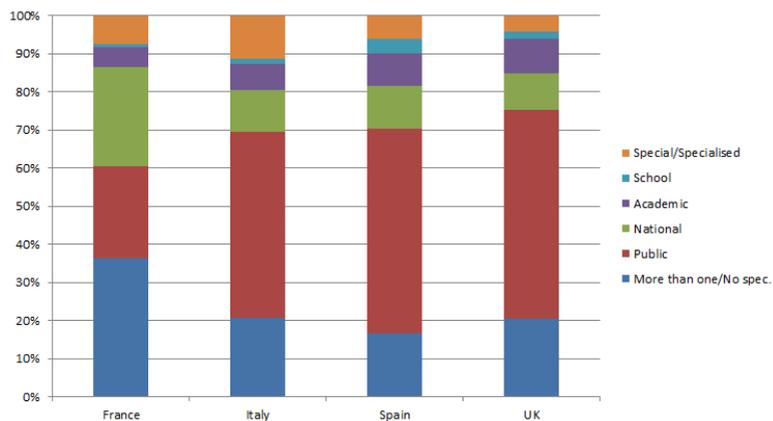
national libraries increases up to 18%. This could mean that, when libraries are only mentioned in writing about more general issues, it is more likely that no specification is given about the type of library considered. The global results concerning the types of libraries deserve a more in-depth analysis in order to verify if there are any differences between the selected countries and newspapers. In the figures 5 and 6 on the following page the data are processed and analysed by country.

From the observation of the two graphs the following considerations emerge:

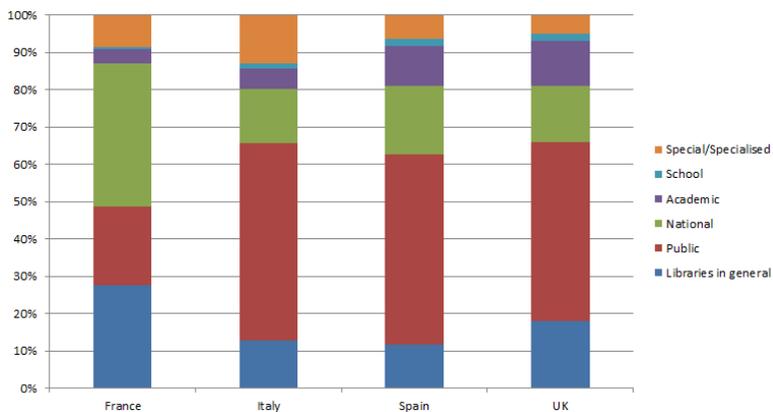
- France is the only country where national libraries, instead of public libraries, are the most present in the debate. This is even truer when considering only those articles mainly talking about libraries; in this case, the number of articles concerning libraries in general declines to the advantage of national libraries;
- the presence of academic libraries in the debate is proportionally higher in the UK and Spain than in France and Italy;
- in Italy the presence of articles on special and specialised libraries is higher than in the other countries and partially balances the low percentage of articles concerning national libraries.

In order to get further elements to interpret the data concerning the most debated types of libraries, in figures 7 and 8 on page 60 these data are processed and analysed by newspaper. From the observations of the two graphs the following considerations emerge:

- as for Spain, France and the UK, the centre-right newspapers (*El Mundo*, *Le Figaro*, *The Times*) seem to give more attention to



**Figure 5:** Types of libraries per country (on total articles)



**Figure 6:** Types of libraries per country (on articles where the library subject prevails)

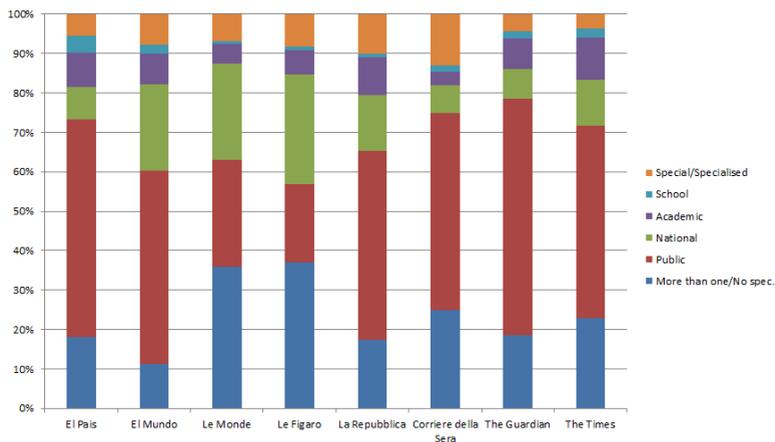
the national libraries in comparison to the centre-left newspapers (*El Pais, Le Monde, The Guardian*), whereas the centre-left newspapers compensate with more articles on public libraries. The only exception is Italy; in this case the number of articles on national libraries is higher in *La Repubblica* than in *Corriere della Sera*, but the centre newspaper has the highest percentage of articles on special and specialised libraries, which can be considered contiguous to the national ones;

- as for the rest, the graphs confirm what emerges from figures 5 and 6 on the facing page.

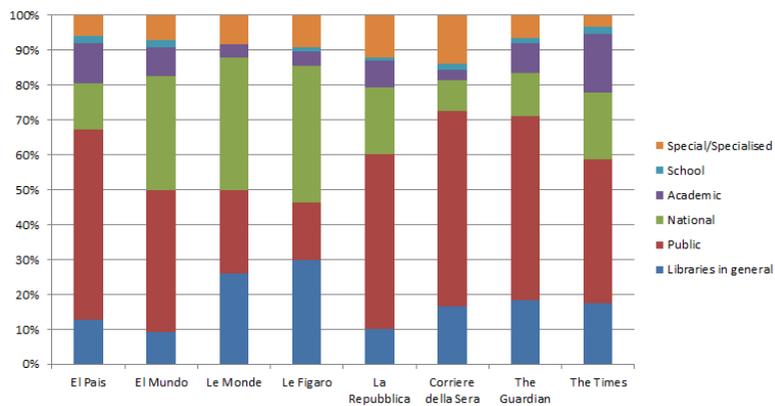
The analysis of the data from the point of view of the topics discussed in the articles offers many insights into the way the European press look at libraries and helps in getting a closer idea of the conception of libraries they tend to convey. First of all, it could be interesting to verify how the topics are distributed on the total number of the articles and on those where the library subject is prevalent (figures 9 and 10 on page 62).

From the analysis of these graphs it is possible to rank the most popular topics concerning libraries (the two percentages in parenthesis refer respectively to the figures 9 and 10 on page 62):

1. Politics/Strategy/Management (28%-22%);
2. Library closures/Budget cuts (17%-14%);
3. Digital/Digitisation (11%-12%)
4. Services/Users (10%-11%)
5. Conservation/Holdings/Catalogue (8%-11%)
6. New libraries/New buildings (7%-9%)
7. Mission/Roles (5%-6%)

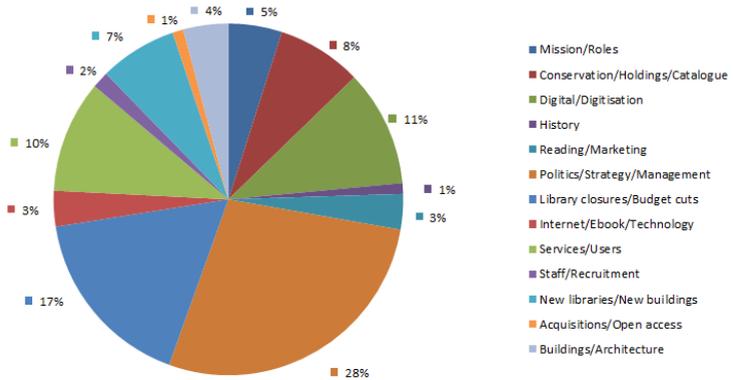


**Figure 7:** Types of libraries per newspaper (on total articles)



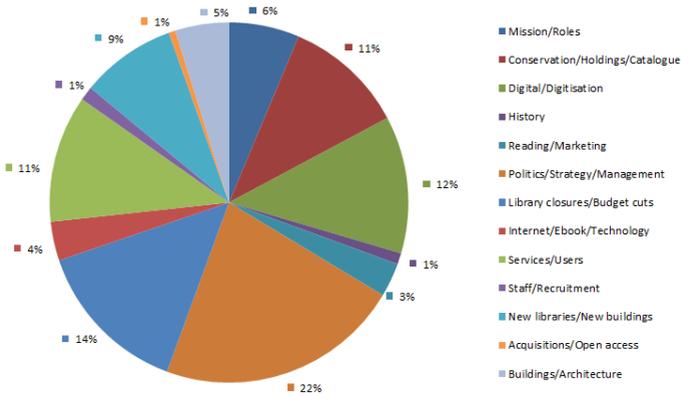
**Figure 8:** Types of libraries per newspaper (on articles where the library subject prevails)

8. Buildings/Architecture (4%-5%)
9. Internet/Ebook/Technology (3%-4%)
10. Reading/Marketing (3%-3%)
11. Staff/Recruitment (2%-1%)
12. History (1%-1%)
13. Acquisition/Open access (1%-1%)



**Figure 9:** Topics (% on total articles)

Considered that newspapers tend to give particular attention to those topics which are of interest for the public opinion, it is quite obvious that the most debated topics concern the least specialised and the most political issues. It should also be highlighted that, considering the overall number of articles, the first two topics cover about 45% of the articles, whereas the percentage declines to 36%



**Figure 10:** Topics (% on articles where the library subject prevails)

if only the articles mainly concerning libraries are taken into account. A possible explanation of this finding is that libraries are frequently mentioned in those articles broadly regarding politics and cuts, while in those articles where libraries are the main subject the contents become more specific. For example, topics like “Conservation/Holdings/Catalogue” and “New libraries/New buildings” are more represented in the articles where the library subject prevails.

Figures 11 and 12 on the next page show the overall distribution of the most popular topics (those beyond the 5% threshold) over the time lapse considered, without distinctions among the four countries, in order to verify which of these topics have been steady and which have been affected by the surrounding political and economic situation over the time.

From the observation of the two graphs a somewhat consistent picture emerges, despite the minor differences between the two. In particular, the following trends are worth being highlighted:

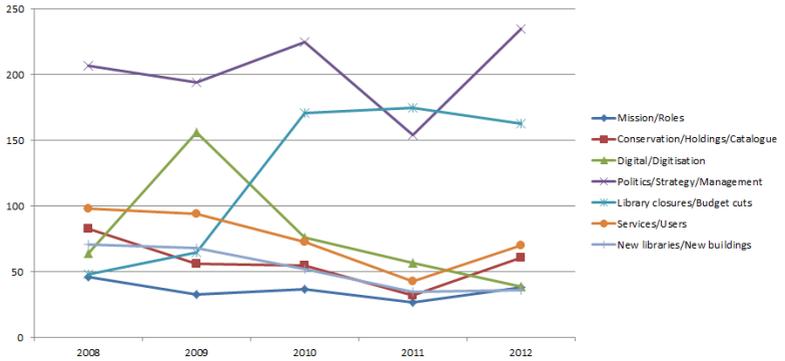


Figure 11: Main topics (over the 5% threshold) per year (on total articles)

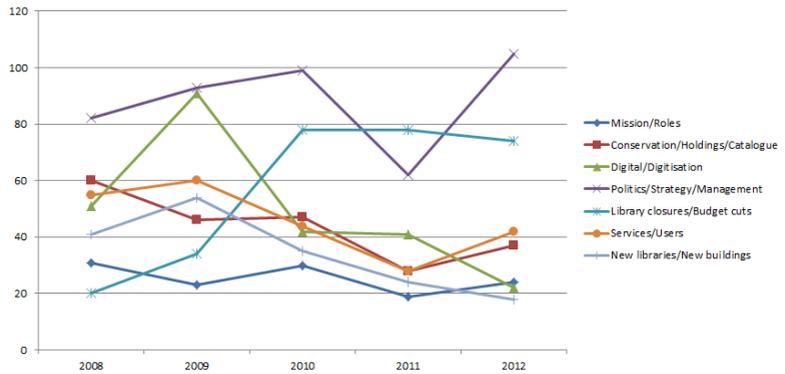


Figure 12: Main topics (over the 5% threshold) per year (on articles where the library subject prevails)

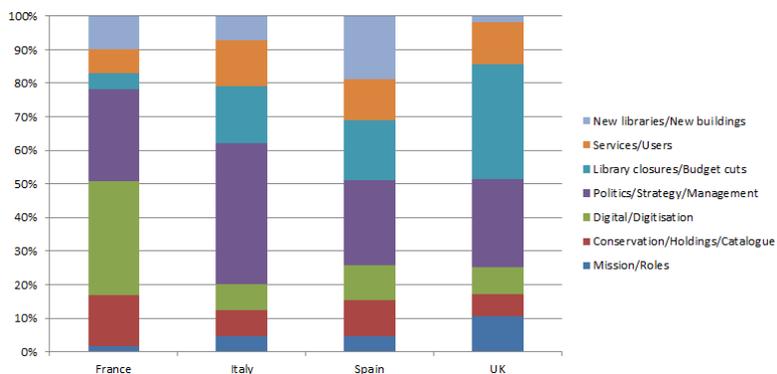
- "Mission/Roles" is the only topic which looks essentially stable during these five years;
- as for "Digital/Digitisation", after a peak in 2009, its presence decreased in the 2010-2012 period; this evolution is probably linked to the rise and decline of the debate after the case involving Google, libraries and publishers on the Google Books project occurred between 2008 and 2009;
- the two topics "Library closures/Budget cuts" and "New libraries/New buildings" have opposite and complementary trends; the first underwent a sharp rise from 2008 to 2010 and kept basically high between 2010 and 2012, whereas the second one had a constant decrease from 2009 on; in this case the link with the economic situation is clear as the ever-increasing negative effects of the crisis have not only stopped the investments, but also brought into questions the survival of many libraries;
- on the contrary, the topics "Services/Users" and "Conservation/Holdings/Catalogue" show fluctuations which seem independent from the external situation; both tend to slightly decline, though from 2011 to 2012 a minor increase was registered;
- the most popular topic "Politics/Strategy/Management" shows a drop in 2011 (which is not easily intelligible), but is basically stable or increasing, as it is a cross-cutting and evergreen issue.

The overall picture emerging from the previous analysis does not take into account the differences between the countries, which could help in further interpreting the trends and better understanding the national specificities and their correlation with the data commented

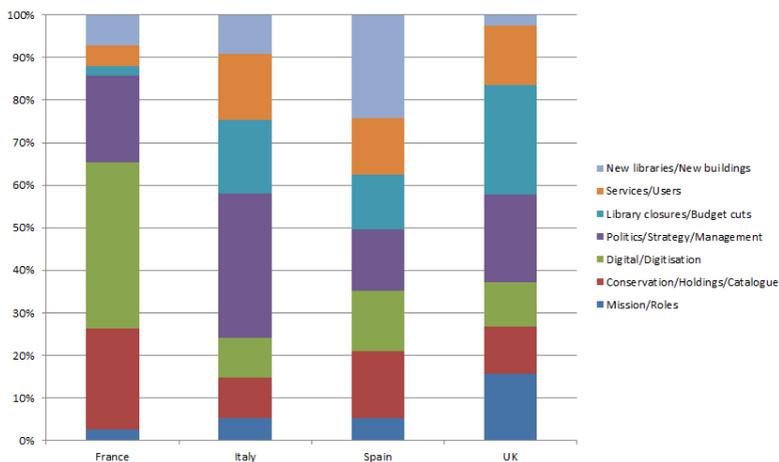
so far. In 13 and 14 on page 66 a comparison among the four chosen countries is proposed concerning the seven most popular topics.

This detailed analysis partly confirms the general trends already illustrated and partly highlights relevant specificities at the national level, which require supplementary considerations. In particular:

- France is the most interested in the "Digital/Digitisation" and "Conservation/Holdings/Catalogue" topics, which are far more present in the French press than in any other country. This finding is consistent with the extensive presence of the national libraries in the French newspapers (see figures 7 and 8 on page 60) and with their traditional high consideration of the national heritage. It is worth remembering that in France digitisation has a central role in the cultural policy and raise important public investments; therefore, the lawsuit concerning the Google Book project - that some French libraries have chosen as partner in digitisation - initiated a lively debate in France between those in favour and those against the private way towards digitisation;
- in Italy the debate on digitisation and digital libraries is relatively scarce which is quite surprising if considering the wide and important printed heritage it can be proud of. After all, despite the immense collections held by the Italian libraries, the national policy concerning digitisation is feeble and highly fragmented; therefore, this issue stays somehow outside the general interest and is debated almost exclusively by librarians and IT people.
- in the UK the "Library closures/Budget cuts" topic seems to dominate the picture, meaning that the debate is lively and deep. Over the last years the news about the hundreds of libraries forced to close or to cut their services due to the



**Figure 13:** Main topics (over the 5% threshold) per country (% on total articles)



**Figure 14:** Main topics (over the 5% threshold) per country (% on articles where the library subject prevails)

governmental austerity policies have crowded the pages of professional reviews and websites, but have also become a matter of general interest flooding the pages of the press;

- as for Spain, the highest percentage of articles concerns new libraries and new buildings, which could be considered quite surprisingly given its well known difficult economic situation. The distribution over time should be verified in order to check if the effects of the economic crisis have changed the terms of the debate during the chronological span considered;
- as far as Italy is concerned, the "Politics/Strategy/Management" topic widely dominates the picture and is far more present than in any other country; when summed with the articles concerning library closures and budget cuts, they cover more than 50% of the most popular topics. The well-known tendency of the Italian politics to be over-represented in the national media seems to be confirmed for the library issues as well.

Some of the findings and questions raised by the above analysis could find an answer by further processing the data. In particular, it could be worth focusing on the two topics which have probably been affected the most by the economic crisis, i.e. "Library closures/Budget cuts" and "New libraries/New buildings", in order to check if the four countries, taken one by one, show different trends over the time lapse considered (figures 15 and 16 on the following page).

As far as the "New libraries/New buildings" topic is concerned, all the countries, with the exception of France, exhibit an evident fall in the number of articles after 2009, though in the UK this trend seems to change after 2010 showing a slight increase. In the case of the Italian newspapers, this finding is complementary to the trend registered for the "Library closures/Budget cuts" topic, which is basically increasing after 2009, though going through ups and

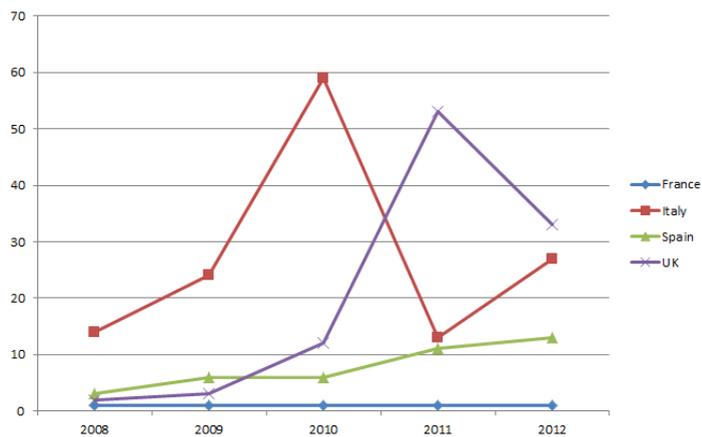


Figure 15: "Library closures/Budget cuts" topic per country (2008-2012)

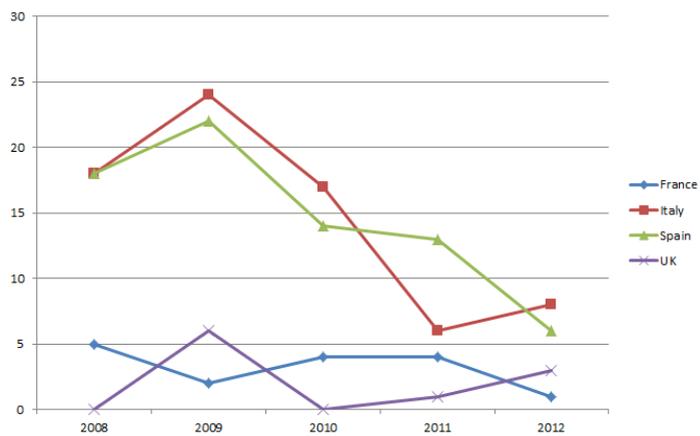
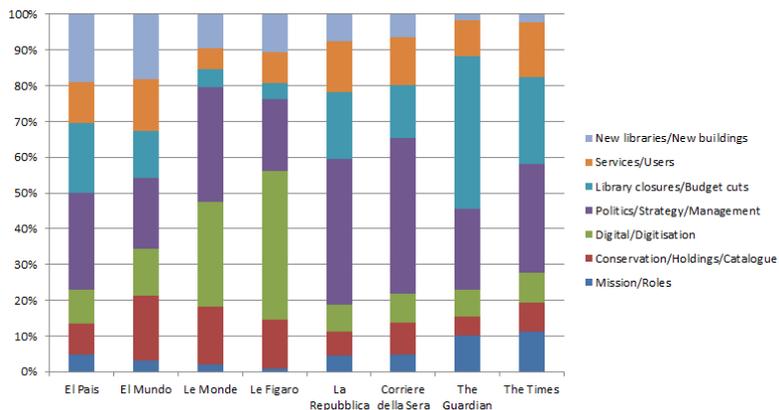


Figure 16: "New libraries/New buildings" topic per country (2008-2012)

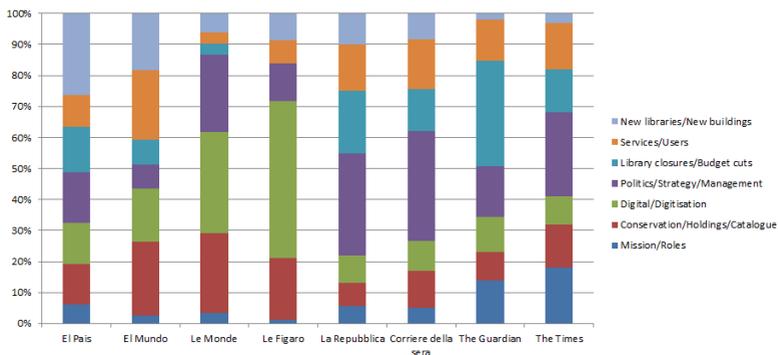
downs, and is in line with the data registered in figures 13 and 14 on page 66 concerning the percent distribution of topics. As for Spain, the significant decline registered for the "New libraries/New buildings" topic is not balanced by the slight increase of the "Library closures/Budget cuts" topic. Given that the Spanish press shows an extensive attention towards new libraries and new buildings, the figures 15 and 16 on the facing page help in getting this topic into the right perspective and reveal that, though central, it has been affected by the crisis in Spain as well as in the other countries, whereas the debate about cuts has become hotter insofar as the economic crisis has deployed its effects. In the UK, while the "New libraries/New buildings" topic keeps low in numbers, though slightly increasing from 2010 on, the debate about library closures and budget cuts seems to burst into flame after 2009, confirming the overall picture coming from the percent distribution of topics. On the contrary, the French newspapers seem not to be involved in the fierce debate about budget cuts and their consequences and register a negligible number of articles on this topic in the whole period, confirming what emerged from figures 13 and 14 on page 66. The articles concerning new libraries and new buildings seem slightly declining, but do not present significant highs and lows. Some more possible keys to interpret the overall picture come from the analysis of the topics distribution in the newspapers. In figures 17 and 18 on the following page a comparison between the eight chosen newspapers is proposed.

The major specificities emerging at newspaper level are the following:

- a comparison between the two graphs highlights that for all the newspapers the number of articles concerning "Politics/Strategy/Management" is higher if considering the total number of articles rather than those where the library subject is



**Figure 17:** Main topics (over the 5% threshold) per newspaper (% on total articles)



**Figure 18:** Main topics (over 5% threshold) per newspaper (% on articles where the library subject prevails)

prevalent. This finding confirms what emerged from figures 9 and 10 on page 62, i.e. when the articles are mainly focused on libraries the issues dealt with tend to be more specific;

- as far as the Spanish newspapers are concerned, there are significant differences among *El País* and *El Mundo*; the first one gives extensive attention to the "New libraries/New buildings" topic and secondly to "Politics/Strategy/Management", whereas *El Mundo* seems more interested in topics like "Conservation/Holdings/Catalogue" and "Services/Users". This could mean that *El País* focuses on more sensitive news for the current political affairs, whereas *El Mundo* has a more conservative and traditional approach;
- looking mainly at Fig. 18, as for the French newspapers, the predominance of the "Digital/Digitisation" topic is true for both newspapers, though *Le Figaro* looks particularly sensitive to this issue which covers more than 50% of the articles concerning the seven most popular topics; in *Le Monde* a relevant number of articles is also devoted to the "Conservation/Holdings/Catalogue" and the "Politics/Strategy/Management" topics;
- the Italian newspapers seem to be more aligned on the most popular topics; *La Repubblica* shows a slightly higher percentage on the "Library closure/Budget cuts" topic, which is balanced by a slightly higher percentage on "Politics/Strategy/Management" registered in *Corriere della Sera*;
- as for the British newspapers, the political positioning seems to have more consequences on the distribution of topics; *The Guardian*, which is the left-wing newspaper, gives a special attention to the most sensitive topic in respect to the governmental politics, i.e. "Library closures/Budget cuts", whereas

*The Times* preferably deals with the more generic debate about "Politics/Strategy/Management" and "Mission/Roles" of libraries;

- an overall analysis shows that no common approaches can be found in those newspapers referring to the same political areas; maybe, a connection between the selection and the choices made by the newspapers and the contingent political situation of the country could be inferred by means of a deeper analysis.

## 6 Conclusions

In this paper the methodology of a research on the European press has been illustrated; this research is aimed at evaluating how the public perception about libraries has evolved over the last five years in four countries: the UK, France, Italy and Spain. The methodology has been explained and discussed to highlight its strengths and weaknesses and some first and temporary findings have been offered in form of graphs and comments. The data analysis has showed the following main trends:

- the prevalence of public and national libraries in the debate, with some significant differences among the four countries;
- a general agreement on the most popular topics, though France seems more interested to the "Digital/Digitisation" topic, whereas the other countries are more sensitive to the political and strategic framework in which libraries are inserted; in general, newspapers confirm to be loudspeakers of topics of general interest, while are less attentive to more specialised issues;
- no clear connection can be traced between the political areas the newspapers belong to and the topics dealt with; a deeper

analysis of the national context is required to better interpret the differences among the data registered in the different newspapers;

- the effect of the economic crisis on the evolution of the debate over the last years is evident both in the presence/absence and in the rise/decline of some topics (e.g. "New libraries/New buildings" declines, while "Library closures/Budget cuts" rises).

As for now, the research proves to be effective not only to highlight the topics and the approaches that newspapers spread in their role of loudspeakers of society, but also to understand how they relate with the common thoughts and the general feelings autonomously emerging from in the society (both influencing them and being influenced by them).

## References

- “America’s Libraries for the 21st Century: An annotated bibliography”. *American Library Association*. (April, 2009). (Cit. on p. 44). Print.
- Barsamian, David and Noam Chomsky. *Propaganda and the public mind: Conversations with Noam Chomsky*. Cambridge: MA.: South East End, 2001. (Cit. on p. 45). Print.
- Bell, David A. “The bookless library”. *New Republic* 2. (2012). (Cit. on p. 44). Print.
- Boldrini, Maurizio. *Il quotidiano: Teorie e tecniche del linguaggio giornalistico*. Milano: Mondadori, 2006. (Cit. on p. 45). Print.
- Bryman, Alan. *Social research methods*. Oxford: Oxford University Press, 2008. (Cit. on pp. 43, 46). Print.
- Chiessi, Sara. “Quanto valgono le biblioteche? Un metodo per valutare l’impatto sociale delle biblioteche pubbliche italiane”. *Bollettino AIB*. (4 2011): 315–327. (Cit. on p. 44). Web.
- Coffman, Steve. “The decline and fall of the library empire”. *Searcher*. (3 2012). (Cit. on p. 44). Web.
- Conti, Sergio. “Ha un futuro la biblioteca pubblica? Spunti e provocazioni (in funzione scaramantica)”. *Bollettino AIB*. (3 2006): 263–267. (Cit. on p. 44). Web.
- Corbetta, Piergiorgio. *La ricerca sociale: metodologia e tecniche*. Bologna: Il Mulino, 2003. (Cit. on p. 43). Print.
- Di Domenico, Giovanni. “Conoscenza, cittadinanza, sviluppo: Appunti sulla biblioteca pubblica come servizio sociale”. *AIB studi*. (1 2013): 13–25. (Cit. on p. 44). Web.
- Faggiolani, Chiara. *Ricerca qualitativa per le biblioteche. Verso la biblioteconomia sociale*. Milano: Editrice Bibliografica, 2012. (Cit. on pp. 43, 46). Print.
- Galluzzi, Anna. *Biblioteche per la città. Nuove prospettive di un servizio pubblico*. Roma: Carocci, 2009. (Cit. on p. 44). Print.
- . “Biblioteche pubbliche tra crisi del welfare e beni comuni della conoscenza. Rischi e opportunità”. *Bibliotime*. (3 2011). (Cit. on p. 44). Print.
- . “Che ne sarà dell’impero bibliotecario?”. *AIB studi*. (3 2012): 363–372. (Cit. on p. 44). Web.
- Giménez Guardiola, Plácido and Manuel Hernández Pedreño. “La imagen social de las bibliotecas en la prensa digital y escrita”. *Anales de Documentación*. (2002): 177–196. (Cit. on p. 45). Web.
- Habermas, Jürgen. *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge: MIT Press, 1989. (Cit. on p. 45). Print.
- Krippendorff, Klaus. *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage, 1980. (Cit. on p. 51). Print.

- Lankes R., David. *The atlas of new librarianship*. Cambridge: MIT, 2011. (Cit. on p. 44). Print.
- Levien, Roger E. *Confronting the future: Strategic visions for the 21st century public library*. ALA, Office for Information Technology Policy, 2011. (Cit. on p. 44). Print.
- Lorusso, Anna Maria and Patrizia Violi. *Semiotica del testo giornalistico*. Roma: Laterza, 2004. (Cit. on p. 45). Print.
- Markless, Sharom and David Streatfield. *Evaluating the impact of your library*. London: Facet, 2006. (Cit. on p. 45). Print.
- Martin, Shannon E. and David Copeland. *The function of newspapers in society: A global perspective*. Westport: Praeger, 2003. (Cit. on p. 45). Print.
- Neuendorf, Kimberly A. and David Copeland. *The content analysis guidebook*. London: Sage, 2002. (Cit. on p. 51). Print.
- Osif, Bonnie A. "W(h)ither libraries? The future of libraries, part 1". *Library Administration & Management*. (1 ). (Cit. on p. 44). Print.
- . "W(h)ither libraries? The future of libraries, part 2". *Library Administration & Management*. (2 ). (Cit. on p. 44). Print.
- Pickard, Alison Jane. *Research methods in information*. Chicago: Neal-Schuman Publishers, 2007. (Cit. on pp. 43, 46). Print.
- Poll, Roswitha. *Bibliography "Impact and outcome of libraries"*. IFLA Publications, 2009. (Cit. on p. 44). Print.
- . "I dati che abbiamo e i dati che potremmo avere: nuovi sviluppi nella valutazione delle biblioteche". *Bollettino AIB*. (4 2011): 369–380. (Cit. on p. 44). Web.
- Salarelli, Alberto. "Baricco, i barbari e la biblioteca pubblica". *JLIS.it*. (1 2011). (Cit. on p. 44). Web.
- Sheridan Burns, Lynette. *Understanding journalism*. London: Sage, 2002. (Cit. on p. 45). Print.
- Solimine, Giovanni. *Gestire il cambiamento. Nuove metodologie per il management della biblioteca*. Milano: Editrice Bibliografica, 2003. (Cit. on p. 43). Print.
- Thelwall, Mike. "Blog searching. The first general-purpose source of retrospective public opinion in the social science?" *Online Information Review*. DOI: [10.1108/14684520710764069](https://doi.org/10.1108/14684520710764069). (3 2007). (Cit. on p. 45). Print.
- Traniello, Paolo. *Biblioteche e società*. Bologna: Il Mulino, 2005. (Cit. on p. 43). Print.
- Ventura, Roberto. *La biblioteca rende: Impatto sociale e economico di un servizio culturale*. Milano: Bibliografica, 2010. (Cit. on p. 44). Print.
- . "La valutazione della biblioteca pubblica: problematiche e strumenti di misurazione dell'impatto culturale, sociale e economico". *Bollettino AIB*. (3 2007): 289–324. (Cit. on p. 44). Web.
- Zelizer, Barbie. *The changing faces of journalism: Tabloidization, technology and truthiness*. London: Routledge, 2009. (Cit. on p. 45). Print.

ANNA GALLUZZI, Biblioteca del Senato "Giovanni Spadolini".

[anna.galluzzi@gmail.com](mailto:anna.galluzzi@gmail.com)

---

Galluzzi, A. "Libraries and public perceptions: A comparative analysis of the European press. Methodological insights". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #8987. DOI: [10.4403/jlis.it-8987](https://doi.org/10.4403/jlis.it-8987). Web.

**ABSTRACT:** Purpose. The paper presents the first results of a research which aims at evaluating how the public perception about libraries has evolved over the last five years in Europe. It quantifies and qualifies the presence of issues concerning libraries in the European press in order to answer the following research questions: which are the most discussed topics and have they changed over the last years? Are there any significant differences between the European countries and newspapers in the debate about libraries? The research is based upon a comparative textual analysis of eight print newspapers of national interest coming from four different European countries (UK, Italy, France and Spain) and concerns the quantity and quality of articles on libraries published from 2008 to 2012. The data analysis highlights the following main trends: public and national libraries are prevalent in the debate; a general agreement is found on the most popular topics; no relevant connection can be traced between the political areas the newspapers belong to and the topics dealt with; a deeper analysis of the national context is required to better interpret the differences among the newspapers; the effect of the economic crisis on the evolution of the debate is evident both in the presence/absence and in the rise/decline of some topics. As for now, the research proves to be effective not only to highlight different topics and approaches, but also to understand how newspapers relate with the common thoughts and the general feelings autonomously emerging from in the society.

**KEYWORDS:** Libraries, European press, Methodology, Content analysis, Public perception

---

Submitted: 2013-05-23

Accepted: 2013-07-05

Published: 2014-01-01





# Dal record al dato. Linked data e ricerca dell'informazione nell'OPAC

Antonella Iacono

## 1 Linked data: la "struttura che connette" il web

«Quale struttura connette il granchio con l'aragosta, l'orchidea con la primula e tutti e quattro con me? E me con tutti voi? E tutti e sei noi con l'ameba da una parte e con lo schizofrenico dall'altra?» (Bateson).

La citazione di Bateson in esergo si riferisce alla necessità, per chi fa ricerca, di allargare gli orizzonti del proprio pensare mettendo in relazione questioni o "cose" apparentemente lontane per ottenere una visione nuova e una comprensione più profonda della realtà.

La sua teoria "ecologica", che lo portava a cercare incessantemente nuove "strutture di collegamento", è un buon esempio per richiamare la necessità di superare la parcellizzazione del sapere contemporaneo e di giungere a nuove forme di apprendimento in grado di cogliere la complessità delle relazioni tra le cose del mondo che ci circonda.

Oggi nel web si assiste a un profondo rinnovamento delle strutture logiche di connessione dell'informazione che si sono evolute



dal collegamento ipertestuale di documenti scarsamente strutturati fino a quello logico delle unità minime di significato, i dati. Il web ha trovato nei linked data una nuova "struttura che connette" l'informazione presente in rete e che rende possibili collegamenti inediti tra i diversi domini del sapere gettando le basi per nuove forme di apprendimento e per una conoscenza interdisciplinare e contestualizzata.

Il termine linked data, nella definizione classica fornita da Tim Berners Lee, fondatore del Web e principale esponente del movimento del Web semantico, si riferisce ad una serie di "buone pratiche" per pubblicare e collegare dati strutturati sul web e formare così la base per il Web semantico.<sup>1</sup>

Linked data significa però, soprattutto, un mutamento di paradigma: vuol dire creare dati che siano "del web" e non solo "sul web" (Coyle) ovvero abbracciare un nuovo modo di creare informazione strutturata che consente ad ogni tipo di informazione di essere strettamente collegata e interoperabile con ogni altra informazione presente nel web.

La prospettiva del Web Semantico coinvolge direttamente lo sviluppo di sistemi di *information retrieval* come i cataloghi elettronici e le modalità di ricerca possibili, consentendo il superamento dell'*information retrieval* tradizionale, facilitando il ritrovamento dei contenuti dei documenti stessi e favorendo, al contempo, l'integrazione con altri documenti presenti nel web con la creazione di rete diffusa di conoscenze. In questo contributo, che si ricollega a quello

---

<sup>1</sup>Com'è noto con i linked data è possibile pubblicare i dati in una modalità leggibile e interpretabile da una macchina, il cui significato è definito da una stringa di parole e marcatori (tripla) per costituire un reticolo di dati collegati appartenenti ad un dominio. Questi sono collegabili ai *dataset* relativi agli altri domini presenti nel web costruendo in tal modo una rete di dati globale, i cui contenuti possono essere scambiati e interpretati dalle macchine, ovvero la base per il Web Semantico (Heath e Bizer; Guerrini e Possemato)

già pubblicato su questa rivista (Iacono), si esplorano le potenzialità e i benefici dell'applicazione dei linked data alla strutturazione logica dei dati bibliografici nel catalogo, o meglio in quello strumento che alla luce dei più recenti sviluppi (Marchitelli e Frigimelica) può essere più propriamente definito un dispositivo per l'accesso e la mediazione dei contenuti della biblioteca, con un'attenzione alle funzionalità di ricerca.

Circoscrivendo la presente osservazione alle potenzialità per la ricerca, il principale vantaggio che può derivare dall'applicazione di questa nuova struttura logica al catalogo risiede infatti nella capacità di contestualizzare l'informazione creando collegamenti tra le conoscenze durante il processo informativo. Il catalogo può diventare lo strumento con cui le biblioteche contemporanee possono colmare le lacune di un sapere oggi sempre più frammentato e agevolare i processi di trasformazione delle informazioni in conoscenza collegando le conoscenze prodotte dalle biblioteche agli altri dati presenti nel Web (Di Domenico).

Con le iniziative intraprese dalle biblioteche nel campo dei linked data ci si avvia verso una nuova importante tappa evolutiva che vedrà verificarsi una radicale trasformazione del catalogo e dei dati bibliografici, del rapporto tra utente e catalogo e tra catalogo e universo globale dell'informazione. Affinché i record bibliografici siano in grado di giocare un rinnovato ruolo nel web è però indispensabile che i dati delle biblioteche non vengano solamente esposti nel nuovo formato Resource Description Framework (RDF), bensì che le basi dati catalografiche adottino RDF come struttura del catalogo (Malmsten), compiendo il fondamentale passaggio dal record al dato.

## 2 Dal record al dato. Il nuovo "record bibliografico collegato" nella gestione del processo informativo

Il *record* bibliografico, principale strumento per rappresentare l'informazione bibliografica nel catalogo, è un insieme di metadati che rappresentano le entità e gli attributi che caratterizzano le dimensioni semiotiche e semantiche di un "oggetto bibliografico" (una monografia, un documento sonoro, etc.) e le relazioni con le altre entità collegate.

Per molto tempo la produzione di record bibliografici si è basata su una struttura piatta e lineare del record, fornita dai formati Machine Readable Cataloguing (MARC), oggi ritenuta obsoleta (Bianchini; Kroeger) e non più adatta a rappresentare compiutamente sia le diverse entità (opera, espressione, manifestazione, item, soggetti, autori, enti, famiglie), sia i legami tra le entità previsti dai modelli IFLA. Più volte l'attuale struttura del record bibliografico è stata vista come il principale ostacolo all'espressione delle entità rappresentate nel catalogo (Yee); un record bibliografico che manca di una granularità adeguata rende difficoltoso lo svolgimento delle basilari funzioni del catalogo (*funzioni-utente*).<sup>2</sup> I linked data offrono la possibilità di operare una profonda ristrutturazione del record bibliografico che si presenta con una nuova struttura granulare e

---

<sup>2</sup>Lo sviluppo del MARC è stato di fatto abbandonato; la Bibliographic Framework Transition Initiative (<http://www.loc.gov/marc/transition>), avviata nell'ottobre 2011 dalla Library of Congress, ha come principale attività lo sviluppo di nuove modalità di creazione e condivisione dei record bibliografici e la ricerca di una soluzione per la sostituzione del formato MARC per lo scambio dei dati bibliografici. Il progetto di un nuovo Bibliographic framework (BIBFRAME, <http://www.loc.gov/bibframe> e <http://bibframe.org>) come nuova modellazione dei dati bibliografici si centra sull'ambiente Web e adotta il data model RDF, i principi e le specifiche dei linked data (Marcum; Kroeger).

atomizzata; il nuovo concetto di record bibliografico si esprime nel "record collegato", costituito virtualmente dalla sintesi di varie triple che combinano dati creati dalle biblioteche e da altri enti.

Attraverso questa rinnovata struttura logica è possibile generare un record formato dalle entità già previste dai modelli concettuali FR dell'IFLA<sup>3</sup> e dotare il record bibliografico di una rete di collegamenti che consentirà all'utente di svolgere efficacemente le funzioni di base del processo informativo.<sup>4</sup> Possiamo, dunque, immaginare il nuovo "record collegato" come prosecuzione dell'idea di "super-record" o "meta record" (Fattahi, *The relevance of cataloguing principles to the online environment; From Information to Knowledge*), recentemente riproposta come una struttura che possa essere agevolmente navigata in modo da fornire all'utente «un ambiente di apprendimento per mezzo del quale il *record* catalografico facilita l'organizzazione e la rappresentazione della conoscenza nel catalogo, come pure la capacità di acquisire, elaborare e interiorizzare la conoscenza da parte di chi utilizza il catalogo». (*From Information to Knowledge*, p. 37).

Il nuovo "record collegato", che si viene a creare quando i dati del catalogo vengono strutturati come linked data, è formato da un collegamento di triple in grado di connettere dati provenienti da diverse fonti sia bibliografiche sia di altra natura, sfruttando un reticolo illimitato di collegamenti tra i dati, il *Web of data*. L'elemento che consente di realizzare questo collegamento è il concetto di Uniform Resource Identifier (URI) ovvero un identificatore unico e persistente associato al dato. Nel nuovo "record collegato" ogni par-

---

<sup>3</sup>Anche lo sviluppo dei modelli FR si sta orientando, non senza difficoltà, verso una rappresentazione ontologica adatta al web semantico; va ricordato in proposito lo sforzo verso la trasformazione di Functional Requirements for Bibliographic Records (FRBR) nella struttura di FRBR-object oriented (FRBRoo), attraverso CIDOC CRM (Le Boeuf, «Foreword»; «A Strange Model Named FRBRoo»).

<sup>4</sup>Per un approfondimento sulle fasi del processo informativo e sulle funzioni che l'utente compie durante l'interazione con l'OPAC si veda Iacono.

te della tripla viene identificata attraverso un URI; avremo dunque URI che identificano le opere, le espressioni, gli autori, i soggetti e le classi, ma anche le relazioni che intercorrono tra le entità saranno a loro volta provviste di URI in modo tale da poter essere riutilizzate. Si crea così un reticolo di dati in cui tutte le entità e le relazioni previste dai modelli FRBR possono essere espresse e nel quale le entità del catalogo possono essere relazionate alle altre entità presenti nel web. Ciò è reso possibile dal recente sviluppo di set di elementi, di vocabolari e ontologie predisposti dalle biblioteche (es. Dublin Core, RDA)<sup>5</sup> o da altre comunità che operano nel Web (Foaf, SKOS). Il collegamento tra i dati di diverso tipo viene creato automaticamente tramite la connessione di diversi dataset con uno sforzo che non ricade unicamente sulle biblioteche, ma viene condiviso tra le biblioteche e le altre comunità del Web Semantico.

### **3 La riduzione delle barriere cognitive della ricerca e il recupero dell'informazione nel contesto.**

Vi sono almeno tre aspetti dell'interazione utente-catalogo che possono beneficiare dell'applicazione dei linked data alla struttura catalografica: la riduzione delle barriere cognitive che intervengono nel processo di ricerca, l'accesso ai dati bibliografici e il miglioramento della ricerca semantica. Nell'interazione che si compie nel catalogo l'utente incontra numerosi ostacoli sia al momento dell'attivazione

---

<sup>5</sup>Le iniziative delle biblioteche nell'ambito del Web Semantico si sono concretizzate nella creazione di dataset che hanno messo a disposizione i dati di cataloghi e bibliografie nazionali in RDF (British National Bibliography, Data.bnf. fr, Biblioteca Nacional de España, Deutsche Nationalbibliografie (DNB) e HBZ Union Catalog, Libris), vocabolari (LCSH, Dewey, LCC, Rameau, Dewey, VIAF) e schemi di metadati (RDA, Dublin Core, FRBR, ISBD etc.).

della ricerca sia lungo il processo informativo (Kulthau; Wilson; Godbold).<sup>6</sup> Tra le barriere maggiormente significative vi sono quelle cognitive (Borgman), che possono consistere:

- in una scarsa comprensione dell'oggetto della ricerca, il focus, presente qualora l'utente si trovi in una fase iniziale della sua ricerca;
- in una difficoltà di espressione dell'oggetto di ricerca, derivante dalla scarsa conoscenza dell'area disciplinare che si sta indagando (domain knowledge);
- nella scarsa conoscenza di un sistema e dei suoi meccanismi di recupero dell'informazione (barriere tecnologiche semantiche e concettuali);
- nella scarsa conoscenza dell'universo bibliografico rappresentato e delle modalità di rappresentazioni dei testi;
- nelle barriere linguistiche.

A queste si possono aggiungere anche quelle barriere provenienti dai modelli mentali che gli utenti hanno costruito nell'utilizzo di altri strumenti di ricerca (Yu).

L'adozione dei linked data come struttura del record bibliografico riduce l'incidenza di questa tipologia di barriere creando le basi per un processo di apprendimento che è fondamentale per il successo di una ricerca nel catalogo. Diviene centrale il concetto di "contesto", concetto che è profondamente connotato nella definizione di conoscenza, in quanto la conoscenza stessa viene creata e

---

<sup>6</sup>Tra le barriere rientrano quelle fisiologiche, cognitive, emozionali, quelle che fanno riferimenti a condizionamenti sociali ed economici (accesso e costo dell'informazione), quelle ambientali e situazionali (tempo a disposizione, fattori geografici, linguistici) e culturali ai quali auspica possa essere riservato un più ampio spazio nella progettazione dei sistemi di ricerca.

appunto collocata in un contesto. Una progettazione efficace del catalogo dovrebbe creare un contesto nel quale l'utente può collocare ogni singola informazione per mezzo della relazione tra le entità costitutive. Il principale ruolo dei linked data nell'innovazione dei cataloghi elettronici risiede quindi nella possibilità che il "web dei dati" venga utilizzato come "contesto esteso" per l'esplorazione dei contenuti che si affianca a quello rappresentato nel catalogo (ossia l'universo delle entità e dei loro collegamenti), per espandere e migliorare il processo di ricerca dell'informazione. I linked data consentono di appoggiare il catalogo ad una infrastruttura cognitiva costituita dalle conoscenze presenti nel "web dei dati", che collega le conoscenze formatesi in contesti diversi.<sup>7</sup>

Tale contesto allargato è in grado di migliorare le seguenti funzioni dell'utente:

**Navigazione e browsing nel catalogo.** L'utente che si trova al momento dell'interazione con l'OPAC nelle prime fasi del processo informativo è naturalmente predisposto a una ricerca esplorativa preferendo modalità attive ma indirette di esplorazione. La tecnologia dei linked data offre varie opportunità, prima fra tutte quella di migliorare i percorsi di navigazione e di approccio al catalogo. Attraverso i linked data è infatti possibile rappresentare le entità catalografiche e consentire una navigazione gerarchica che dal raggruppamento delle opere si svolga fino all'item. In tal modo si compiono le "funzioni bibliografiche" che l'utente svolge nel catalogo e si realizza una navigazione significativa per l'utente e, al contempo, un processo di apprendimento. La navigazione attraverso il colle-

---

<sup>7</sup>Lo schema di un reticolo bibliografico espresso con i linked data nell'OPAC Libris della Biblioteca nazionale svedese è visibile all'URL: <http://librisbloggen.kb.se/2008/12/03/libris-available-as-linked-data> Un data model è quello della British Library visibile in <http://www.bl.uk/bibliographic/pdfs/bldatamodelbook.pdf>.

gamento ad altri *dataset* (es. DBPedia) può estendersi al web, inglobando dinamicamente informazioni provenienti da fonti esterne.

**Interrogazione tramite formulazione di una query.** Nelle fase intermedia del processo di ricerca, in cui l'utente è in grado di formulare interrogazioni al sistema, è invece estremamente importante offrire un aiuto nel formulare *query*. Sebbene gli OPAC oggi ricorrano all'autocompletamento, sovente questa tecnica non si appoggia a vocabolari controllati, rivelandosi inefficace. I linked data, che cominciano ad avere alcune prime applicazioni nello sviluppo dei cataloghi elettronici (Malmsten).<sup>8</sup> Vi sono anche prime esperienze di integrazione in discovery tools open source come VuFind (Hatop), consentono di usufruire di ontologie già create in maniera dinamica e automatica all'interno del Web Semantico e dunque utilizzabili a vantaggio della ricerca nei cataloghi per aiutare l'utente a formulare interrogazioni. Tramite i legami tra diversi *dataset*, si offre all'OPAC una struttura sindetica potenzialmente illimitata che può servire da vocabolario d'accesso al catalogo. Ad esempio, gli accessi semantici di un catalogo (come le stringhe LCSH o le notazioni delle classi) vengono automaticamente mappati ai concetti correlati e diventano recuperabili non solo a partire dai termini che compongono il soggetto, ma anche a partire da termini semanticamente simili, perché collegati allo stesso concetto. La rete dei dati presente nel Web dunque potrebbe utilmente essere sfruttata dalle biblioteche per sviluppare interfacce che aiutino l'utente a scegliere il soggetto più appropriato per la propria ricerca. La tecnologia dei linked data consente anche di abbattere le barriere linguistiche all'utilizzo di un catalogo, in quanto tutte le forme varianti

---

<sup>8</sup>Un esempio è l'OPAC Libris <http://libris.kb.se>.

di un nome, di un titolo o di un soggetto possono essere collegate alla forma accettata secondo il meccanismo già noto dell'*authority file* con un vantaggio in più per la ricerca semantica: la possibilità di esprimere un concetto in una lingua diversa da quella per cui il catalogo è progettato; i termini dell'interrogazione sebbene in una lingua diversa da quella espressa in un catalogo possono venire ugualmente collegati automaticamente tramite i linked data con un indubbio vantaggio per chi ricerca. Pensiamo alla possibilità di collegare i dati degli archivi di autorità: collegando il dato alla forma in VIAF,<sup>9</sup> l'utente potrebbe trovare una descrizione catalografica partendo dalla forma desiderata tramite il collegamento a tutte le forme varianti. Allo stesso modo ogni dato può essere collegato o provenire da archivi di diverso tipo: statistici, di anagrafi pubbliche, di enti territoriali o governativi, etc.<sup>10</sup> Un ulteriore beneficio è quello di poter interrogare il catalogo servendosi del linguaggio naturale per essere indirizzati ai termini più adatti per la ricerca nel catalogo o lanciare una ricerca esplorativa a partire dai termini suggeriti dal sistema.

**Valutazione dell'informazione ricevuta.** La valutazione dei risultati di un'interrogazione o dei record recuperati nel corso della ricerca nel catalogo consiste, per l'utente che si trova nelle prime fasi del processo di ricerca, essenzialmente nel valutare la rispondenza all'argomento ricercato (rilevanza topica), mentre per chi si trova già nelle fasi successive anche alla corrispondenza soggettiva, ovvero un'ultima analisi alla pertinenza e

---

<sup>9</sup><http://www.viaf.org>.

<sup>10</sup>Tra i progetti è da citare l'arricchimento di degli articoli di Wikipedia con i record d'autorità VIAF da parte dei wikipediani di OCLC e British library (Klein e Alex), esperimento che viene replicato nel nostro paese dai bibliotecari wikipediani con i gli authority records SBN; l'operazione sarebbe ancor più interessante se ampliata alle voci d'autorità per i soggetti.

all'utilità. L'opportunità offerta dalla tecnologia dei linked data tramite il record bibliografico "collegato" risiede nella possibilità di ampliare il contesto utile all'utente per effettuare la valutazione della rilevanza e della pertinenza, tramite il supporto delle migliaia di risorse presenti nel Web Semantico. La quantità di dati recuperabile dinamicamente attraverso un record collegato al *Web of data* è non solo potenzialmente illimitata, ma arricchisce il risultato di informazioni supplementari sempre aggiornate, migliorando la valutazione: ad esempio la possibilità di visualizzare un raggruppamento di tutti gli argomenti trattati da un autore nelle sue opere, di visualizzare la fortuna editoriale di un autore in un dato periodo di tempo o l'autorevolezza, sono funzionalità tutto sommato non particolarmente complesse, ma di grande aiuto all'utente, che possono essere implementate per mezzo del collegamento del record ai dati presenti nel Web.<sup>11</sup>

**Accesso alle risorse bibliografiche.** La novità risiede nella possibilità di accedere al catalogo da fonti esterne; una volta che i dati bibliografici entrano a far parte del Web of data essi saranno oggetto del recupero dell'informazione da qualsiasi altro dato collegato al *Web Semantico*.

**Visualizzazioni avanzate.** Una delle novità introdotte dagli OPAC di nuova generazione è certamente la componente visuale del recupero dell'informazione, manifestatasi non solo nella

---

<sup>11</sup>In questa direzione un'iniziativa interessante è Social Networks and Archival Context project (SNAC), un sistema per la ricerca storica a partire dai record archivistici di enti, persone e famiglie che utilizza i linked data per ricostruire il contesto storico e sociale (persone, enti date, luoghi, eventi, soggetti correlati, cronologie) e del quale un prototipo è visibile in <http://socialarchive.iath.virginia.edu/index.html>. Dei collegamenti al *Web semantico* si serve oggi anche Google per presentare accanto alla SERP i risultati provenienti estratti da Wikipedia.

comparsa delle copertine dei volumi quale arricchimento del record, ma soprattutto nella predisposizione di nuove modalità di browsing visuale delle collezioni e nella costruzione di canali di esplorazione delle risorse. Non c'è dubbio che oggi molte applicazioni del web si distinguano anche per la loro veste accattivante che si esprime, sovente, nella creazione di mappe visive. Il record bibliografico 'collegato', che la logica dei linked data consente di costruire, getta le basi per la costruzione di «ambienti arricchiti di conoscenza» (Fattahi, *From Information to Knowledge*, p.36) nei quali sarà possibile per l'utente attivare nuovi percorsi di esplorazione del catalogo a partire dalle entità oggetto di interesse. Non dobbiamo infatti dimenticare che le nuove possibilità di implementazione di FRBR ai cataloghi renderanno possibile in futuro organizzare i risultati delle ricerche tramite raggruppamenti logici previsti dal modello. Il "record collegato" compie un ulteriore passo in avanti: attraverso il collegamento dei dataset delle biblioteche con quelli presenti nel *Web dei dati* si può dare luogo alla realizzazione di rappresentazioni ontologiche dei risultati delle ricerche che gli utenti mettono in atto, capaci di aprire orizzonti inaspettati per la ricerca semantica.

## **4 I Linked data e i nuovi orizzonti della ricerca semantica**

Il *data model* fornito da RDF, che è alla base dell'architettura dei linked data, consente di strutturare l'informazione bibliografica in modo da potenziare le funzionalità della ricerca semantica. Ciò è possibile grazie al fatto che nel Web Semantico tutti gli URI che si riferiscono allo stesso concetto sono correlati creando una vasta rete

semantica di relazioni che può essere automaticamente collegata al catalogo. Come è noto, i cataloghi soffrono della contingenza e della difficoltà per l'utente dovuta all'utilizzo dei dispositivi indicali: essi rappresentano una necessità, ma pongono all'utente il problema di misurarsi con il "linguaggio bibliografico" del catalogo (Serrai, *Indici, logica e linguaggio: problemi di catalogazione semantica; Del catalogo alfabetico per soggetti: semantica del rapporto indicale*; Ventura). Per questo motivo gli accessi semantici sono oggi sempre meno utilizzati dagli utenti che preferiscono immettere parole chiave ed effettuare una ricerca *Google like* su tutti i campi ottenendo sovente risultati imprecisi.

Gli indici che le biblioteche utilizzano ai fini dell'organizzazione delle registrazioni catalografiche e del loro recupero riflettono tuttavia il modo in cui le biblioteche organizzano il sapere. Per tradizione esse hanno seguito strategie di indicizzazione *content oriented* ovvero tese ad individuare con precisione nei termini indicali il contenuto di un documento in base all'utenza di riferimento, sacrificando così le potenzialità di utilizzazione dei libri stessi in quanto le molteplici dimensioni e interpretazioni che possono provenire da studiosi di formazione e provenienza scientifica diversa vengono ridotte alla creazione di un'unica stringa coestesa (Biagetti).<sup>12</sup>

---

<sup>12</sup>Biagetti analizza il ruolo degli indici semantici nel contributo *Indici semantici e ricerca scientifica* che, ad una rilettura alla luce delle nuove tecnologie del web semantico, offre interessanti spunti di riflessione. La studiosa rileva come rispetto agli aspetti sintattici del processo di indicizzazione sia stata tradizionalmente riservata una minore attenzione all'analisi dei contenuti semantici e alla scelta delle strategie di indicizzazione da adottare. Ciò è dovuto – afferma la studiosa – al prevalere di metodologie che si sono richiamate al procedimento della coestensione ovvero alla necessità di condensare in una sola stringa i molteplici contenuti semantici di un libro. A tale strategia orientata al contenuto *content oriented* si contrappone, specie nell'indicizzazione di documenti che ricadono nelle Scienze umane, la possibilità di studiare strategie *need-oriented*, ovvero indirizzate alle necessità e alle possibili richieste di ricercatori formati in contesti scientifici diversi (, p.16 «Indici semantici e ricerca scientifica»).

Oggi i Linked data ci offrono l'opportunità di collegare le sistemazioni concettuali messe in opera dalle biblioteche alle altre modalità di organizzare e rappresentare le conoscenze create da altre comunità che operano nel web:

«An essential process is the joining together of subcultures when a wider common language is needed. Often two groups independently develop very similar concepts, and describing the relation between them brings great benefits. . . The Semantic Web, in naming every concept simply by a URI, lets anyone express new concepts that they invent with minimal effort. Its unifying logical language will enable these concepts to be progressively linked into a universal Web» (Berners-Lee, Hendle e Lassila).

L'affermazione di Berners Lee, Hendler e Lassilla esprime l'enorme potenzialità dell'applicazione del Web Semantico ai cataloghi: il fatto di essere una logica unificante che consente di collegare i diversi linguaggi che nel web vengono utilizzati per denominare cose, eventi, luoghi, concetti, persone. Le principali potenzialità applicative che i linked data offrono alla ricerca semantica sono ravvisabili in almeno due principali punti:

- mappatura di concetti presenti in schemi e sistemi di organizzazione delle conoscenze diversi ai fini dell'interrogazione, della visualizzazione e della navigazione nel catalogo;
- sviluppo di modalità multidimensionali per la ricerca.

Il Web semantico fornisce un'infrastruttura tecnologica che consente di collegare allo stesso concetto identificato da un URI i termini o le espressioni utilizzate da diverse comunità del Web ottenendo una rete sindetica già pronta che le biblioteche possono utilizzare per i

propri scopi. La scomposizione dei soggetti e delle classi nei concetti che li compongono potrà unire diversi sistemi di organizzazione delle conoscenze (KOS) creati dalle biblioteche e sistemi creati da altre comunità del web semantico, offrendo modalità avanzate di recupero dell'informazione che si possono avvalere di una rete di rimandi tra termini già pronta all'uso. Sarà dunque possibile raggiungere un soggetto o una classe a partire da termini provenienti da KOS di diverso tipo e collegare tra loro termini affini e concetti simili ampliando il raggio della ricerca semantica con indubbi vantaggi per l'utente. Il *data model* fornito da RDF è molto simile a quello delineato da FRSAD, lo schema concettuale di IFLA per modellare la circolarità delle opere (*aboutness*) che adotta un modello fondato sulle entità *thema* (concetto) e *nomen* (etichetta).<sup>13</sup> FRSAD trova adeguate corrispondenze nei due schemi di codifica per il Web semantico: Simple Knowledge Organization System (SKOS) e Web Ontology Language (OWL),<sup>14</sup> che forniscono modelli per esprimere le strutture di sistemi di organizzazione delle conoscenze.

In particolare, SKOS definisce classi e proprietà per esprimere la struttura di base e il contenuto di schemi concettuali come thesauri, schemi di classificazione, tassonomie, soggettari, folksonomies e altri vocabolari controllati. Questo standard adotta un modello basato sul "concetto" e sulla separazione tra concetti e etichette (ovvero i termini per esprimerli) già espressa in FRSAD; il suo vocabolario è modellato sui concetti che a loro volta sono rappresentati da etichette e che sono confrontabili con le entità definite in FRSAD ovvero *thema*, *nomen* e i loro attributi. Attraverso SKOS i concetti possono essere organizzati in gerarchie e reti di associazioni, aggregati in schemi

---

<sup>13</sup>In particolare si veda IFLA, FRSAD Final Report: Appendix C: FRSAD Model and Other Models, p. 48-50.

<sup>14</sup>Rispettivamente <http://www.w3.org/2009/08/skos-reference/skos.html#ConceptScheme> e [http://www.w3.org/TR/#tr\\_OWL\\_Web\\_Ontology\\_Language](http://www.w3.org/TR/#tr_OWL_Web_Ontology_Language)

concettuali, ordinati in collezioni ordinate e mappati ai concetti presenti in altri schemi (Žumer e Zeng).

Il modello definito da FRISAD trova mappature appropriate nelle classi, negli attributi e nelle relazioni semantiche stabilite dallo standard del W3C SKOS: frsad:thema equivale a skos:Concept; frsad:nomen equivale a skos:label e altre proprietà come altre proprietà come skos:broader, skos:narrower, skos:related consentono di esprimere le relazioni gerarchiche e associative tra i concetti tipiche di un'organizzazione thesaurale. Nella costruzione di una tripla ogni concetto viene associato – tramite il predicato `rdf:type` – alla classe specifica `skos:Concept` ed identificato in modo univoco da un URI. Lo standard prevede la possibilità di esprimere un'etichetta in diverse lingue, di identificare i termini che all'interno dello schema concettuale sono accettati e usati come descrittori nel sistema d'indicizzazione, e infine, di descrivere anche gruppi di termini che sono di corredo al concetto e che non hanno un'identità semantica e che quindi non verranno identificati con un URI, ma solo con etichette (es. i sinonimi o i quasi sinonimi le varianti ortografiche, le abbreviazioni e gli acronimi).<sup>15</sup> Web Ontology Language (OWL) è un altro linguaggio per esprimere ontologie di concetti nel Web semantico con un significato formalmente definito. Esso permette di esplicitare le relazioni tra vari concetti (*thema*) previsti da FRISAD in maniera più completa rispetto a SKOS in quanto consente di esprimere ulteriori relazioni tra le classi come `SubClassOf`, `EquivalentClasses`, `DisjointClasses`, e `DisjointUnion`. I linguaggi SKOS e OWL consentono in combinazione di esprimere "concetti" ovvero accessi per

---

<sup>15</sup>Questo tipo di relazioni tra etichette possono essere risolte ricorrendo a un'apposita estensione di SKOS, SKOS-XL, che fornisce un supporto per descrivere entità lessicali (termini) attribuiti ai concetti. Lo scopo è reificare le etichette di `skos:Concepts`, trattandole come risorse RDF effettive. Ciò consente la possibilità per esse di essere ulteriormente annotate utilizzando ad esempio una proprietà "è una traduzione di". Lo schema è visibile in: [xmlns:xl=http://www.w3.org/TR/skos-reference/skos-xl.html](http://www.w3.org/TR/skos-reference/skos-xl.html)

soggetti e per classi e le loro relazioni, ma anche di mappare concetti presenti in diversi schemi.

L'assegnazione di un identificatore univoco ai concetti facilita l'interoperabilità tra KOS differenti, cioè la possibilità di mappare entità semantiche facenti parte di schemi concettuali diversi. Per stabilire questi collegamenti lo standard SKOS definisce alcune particolari proprietà che prevedono differenti livelli di equivalenza (*skos:closeMatch*; *skos:exactMatch*; *skos:broaderMatch* e *skos:narrowerMatch*; *skos:relatedMatch*). Di recente, i più importanti e diffusi soggettari e schemi di classificazione sono stati tradotti nella forma di Linked data, tra questi i vocabolari della Library of Congress (LCSH, LCC), la Classificazione Decimale Dewey, RAMEAU e il Thesaurus del Nuovo Soggettario consentendo questa tipologia di collegamenti. Sarà così possibile collegare schemi diversi di organizzazione delle conoscenze (KOS) creati dalle biblioteche come soggetti e classi<sup>16</sup> o schemi completamente diversi per tipologia e profondità come un soggettario e una folksonomia offrendo nell'OPAC il collegamento automatico dei record di autorità a varie altre forme presenti sul web che possono diventare, se collegate ai dati catalografici, punti d'accesso aggiuntivi al catalogo. Attraverso il collegamento e la mappatura di concetti presenti in differenti schemi la ricerca per soggetto può essere enormemente potenziata, sia nella fase di interrogazione, in quanto è ormai possibile utilizzare la rete di collegamenti del record per aiutare l'utente a formulare interrogazioni, sia in fase di navigazione in quanto è possibile produrre visualizzazioni avanzate che sfruttano il collegamento tra schemi

---

<sup>16</sup>Questo tipo di collegamento, che non viene ancora implementato negli OPAC, è un espediente che favorisce la navigazione e l'apprendimento che si può realizzare durante il processo informativo. L'utente può infatti partire da un soggetto per visualizzare tutte le classi che utilizzano quel soggetto e navigare nell'albero classificatorio correlato, o al contrario visualizzare tutti i soggetti che sono legati alle opere che fanno parte di una determinata classe

utilizzati sia nel settore bibliotecario che in altri settori scientifici, per espandere ulteriormente la ricerca. È inoltre possibile offrire all'utente l'opportunità di interrogare il catalogo secondo una logica sfumata: la possibilità di collegare concetti non identici ma simili, prevista nel Web Semantico, rende realizzabile il recupero di un concetto partendo da concetti ad esso correlati. In questo modo viene migliorata la ricerca semantica e la fruizione dei documenti stessi, che potranno essere più facilmente recuperati e utilizzati.

Queste nuove possibilità aprono prospettive inedite per la ricerca semantica sia in ambito generalista che in ambito scientifico: se le biblioteche, in base alla conoscenza della propria utenza di riferimento, hanno adottato per molto tempo strategie di indicizzazione semantica rivolte alla sommarizzazione estraendo dal documento i contenuti per ricavarne una circolarità (*aboutness*) adatta ai lettori di riferimento, oggi invece, nei cataloghi elettronici la cui utenza è remota e non conoscibile a priori, si registra una mediazione minima da parte del catalogatore (Biagetti). Questa pratica – affermava giustamente la studiosa – contrasta con la visione pragmatista del soggetto espressa da Hjørland nella teoria della domain analysis (analisi di dominio) secondo la quale le proprietà che distinguono semanticamente un documento possono avere significati diversi all'interno di diversi ambiti disciplinari e che richiederebbe l'applicazione di modalità di indicizzazione *user oriented*.

La teoria del soggetto come «potenziale epistemologico dei documenti» elaborata dallo studioso danese (citato in Biagetti), afferma che le proprietà significative del soggetto non possono essere stabilite *a priori*, ma dipendono dal contesto nel quale vengono elaborate, ovvero dalla società stessa e che lo sviluppo epistemologico di un soggetto è frutto dell'avanzamento scientifico e sociale. In questo senso emergono gli elementi di novità e le nuove opportunità che la tecnologia dei linked data offre alla rielaborazione di una rinnovata

“teoria del soggetto”:<sup>17</sup> il passaggio da una rete di “link” ad una rete di “significati” che si ottiene con la creazione del Web semantico, consente di allargare il raggio di analisi alle modalità di rappresentazione della conoscenza in diverse comunità scientifiche ed equivale alla possibilità per i nuovi sistemi di recupero dell’informazione di usufruire di quella struttura di significati che costituisce il portato epistemologico di un soggetto, creando le condizioni per proficui scambi disciplinari e per un recupero dell’informazione qualitativamente differente. Tornando alla ricerca nell’OPAC, particolarmente interessanti si rivelano i salti dimensionali che l’utente può compiere, che introducono nella ricerca semantica funzionalità decisamente innovative rese possibili dal collegamento di *data set* bibliografici a dataset di altra natura (ad esempio, quelli di persone, geografici, o tematici es. di film, di opere teatrali) o di altre istituzioni del patrimonio culturale come musei e archivi. Tramite il collegamento con i *data set* del patrimonio culturale sarà possibile per l’utente saltare da un soggetto ad un oggetto museale che lo rappresenta, (ad esempio un quadro, una scultura). Sarà, infine, possibile attivare collegamenti tra dati bibliografici alle “cose” reali descritte nel web; ad esempio un utente, attraverso il collegamento tra *data set* delle biblioteche e database geografici, potrà visualizzare in una mappa geografica i soggetti correlati o gli autori che pubblicano in un determinato paese o che hanno avuto più influenza in una certa area geografica. Un’ulteriore possibilità offerta dalle tecnologie del web semantico riguarda la progettazione delle interfacce dell’OPAC: la visualizzazione dei dati bibliografici potrà avvalersi delle tecnologie del Web semantico che consentono, ad esempio, la

---

<sup>17</sup>Così si esprimeva in proposito Biagetti (p. 24): «La Biblioteconomia e la Scienza dell’informazione, secondo l’impostazione offerta da Birger Hjørland, devono elaborare una “Teoria del soggetto” che permetta di produrre rappresentazioni corrette della conoscenza espressa nei libri, anche prendendo in considerazione l’ampio orizzonte costituito dai diversi campi scientifici».

realizzazione di pagine dinamiche per ogni entità di interesse del catalogo, ad esempio un'opera, un soggetto, un autore. In tal modo viene ampliata la possibilità di recupero di un'entità nel suo contesto rappresentando la sua collocazione nell'universo bibliografico del catalogo e usufruendo di collegamenti con gli altri dati nel Web. Un esempio di come potrebbe apparire la pagina di un soggetto è quella fornita dal progetto della Bibliothèque nationale de France<sup>18</sup> (Wenz): a partire da un determinato soggetto viene creata dinamicamente una pagina che offre collegamenti a documenti, film, immagini, risorse multimediali e collegamenti esterni. Visualizzazioni avanzate create a partire dalle entità di interesse del catalogo consentono il recupero dell'informazione bibliografica in un contesto arricchito da altre risorse autorevoli provenienti dal Web. Si passa così da un recupero ispirato all'*information retrieval* ad un recupero in base al modello entità-relazione, che si rivela particolarmente efficace nel raggruppamento dei concetti nella ricerca semantica.

## 5 Le nuove basi dell'accesso bibliografico

Ciò che, infine, segnerà profondamente il processo di evoluzione del catalogo elettronico e il rapporto utente-catalogo sarà lo spostamento a livello del Web dei meccanismi del controllo bibliografico e dell'organizzazione della conoscenza, ovvero delle principali vie d'accesso al catalogo. I linked data aprono nuove prospettive di accesso all'informazione bibliografica allo scopo di garantirne la fruibilità e il riutilizzo gettando nuova luce sui temi dell'interoperabilità, della pluralità e della convergenza che da decenni sono oggetto di riflessione e che sono alla base di un'adeguata progettazione di ambienti digitali.

---

<sup>18</sup><http://data.bnf.fr>.

A questo punto l'OPAC si trasformerà da strumento autoreferenziale a strumento profondamente integrato nel web e in altri ambienti digitali, realizzando quelle condizioni di interoperabilità da sempre auspiccate favorendo così la ricerca e l'uso dei dati bibliografici da chiunque cerchi informazioni nel web e non più dai soli utenti del catalogo (Baker).

Le tecnologie dei linked data potranno facilitare il raggiungimento delle quattro principali tipologie di interoperabilità che consentono la costruzione di ambienti digitali integrati (Leombroni) e cioè quelle dell'interoperabilità tecnologica, semantica, organizzativa e delle risorse umane:

1. *Interoperabilità tecnica*: è la garanzia dell'interoperabilità fra sistemi e applicazioni; essa è resa possibile dall'adozione di una tecnologia unificante come quella fornita dal data model RDF comune a tutto il web;
2. *Interoperabilità semantica*: la tecnologia dei linked data consente alle varie comunità coinvolte di definire concetti il cui significato è condiviso tramite il meccanismo degli URI che collega ai concetti i termini con i quali le diverse comunità li identificano e utilizzano per la ricerca;
3. *Interoperabilità delle risorse umane*: i linked data permettono una convergenza dei dati e dei servizi che attorno ad essi si sviluppano; a tal fine favoriscono l'interoperabilità delle competenze distribuite nel Web Semantico che vengono collegate insieme per lo sviluppo di nuovi servizi basati sui dati;
4. *Interoperabilità organizzativa*: i linked data consentono di sviluppare una convergenza dei servizi delle istituzioni che operano nell'ambito del cultural heritage, della ricerca scientifica, dei servizi culturali, che hanno il delicato compito di cooperare

con le altre comunità del Web affinché venga garantita la qualità e l'integrità e dei dati immessi attraverso idonee procedure di certificazione della provenienza dei dati (Crupi; Morando). Quest'ultima tipologia comprende anche una nuova interoperabilità giuridica relativa alle licenze con le quali i dati vengono pubblicati nel *Web Semantico*, raggiungibile attraverso l'armonizzazione delle licenze d'uso presenti a livello internazionale tramite l'adozione di licenze di utilizzo dei dati più appropriate, e possibilmente, aperte.

## Riferimenti bibliografici

- Baker, Thomas. «Designing data for the open world of the Web». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-6308](https://doi.org/10.4403/jlis.it-6308). (2013): 63–66. (Cit. a p. 97).
- Bateson, Gregory. *Mind and Nature: A Necessary Unity*. New York: Dutton, 1984. (Cit. a p. 77).
- Berners-Lee, Tim, James Hendle e Ora Lassila. «» (). (Cit. a p. 90).
- Biagetti, Maria Teresa. «Indici semantici e ricerca scientifica». *L'organizzazione del sapere. Studi in onore di Alfredo Serrai*. A cura di Maria Teresa Biagetti. Milano: Sylvestre Bonnard, 2004. 15–29. (Cit. alle pp. 89, 94, 95).
- Bianchini, Carlo. «Dagli OPAC ai library linked data: come cambiano le risposte ai bisogni degli utenti». *AIB Studi* 52.3. DOI: [10.2426/aibstudi-8597](https://doi.org/10.2426/aibstudi-8597). (2012): 303–323. (Cit. a p. 80).
- Borgman, Christine L. «Why are online catalogs still hard to use?» *Journal of the American Society for Information Science* 47.7. (1996): 493–503. (Cit. a p. 83).
- Coyle, Karen. «Linked Data: an Evolution». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-5443](https://doi.org/10.4403/jlis.it-5443). (2013): 53–62. (Cit. a p. 78).
- Crupi, Gianfranco. «Beyond the Pillars of Hercules: Linked data and Cultural heritage». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-8587](https://doi.org/10.4403/jlis.it-8587). (2013): 25–49. (Cit. a p. 98).
- Di Domenico, Giovanni. «Conoscenza, cittadinanza, sviluppo: appunti sulla biblioteca pubblica come servizio sociale». *AIBS Studi* 53.1. (2013): 13–25. <<http://aibstudi.aib.it/article/view/8875>>. (Cit. a p. 79).
- Fattahi, Rahmatollah. *From Information to Knowledge: SuperWorks and the Challenges in the Organization and Representation of the Bibliographic Universe: Lectio magistralis in Biblioteconomia*. Firenze: Casalini Libri, 2010. (Cit. alle pp. 81, 88).
- . *The relevance of cataloguing principles to the online environment: an historical and analytical study: a thesis submitted in fulfillment of the requirements for the degree of doctor of philosophy in the School of information, library and archive studies*. Sidney: University of South Wales, 1997. <http://profsite.um.ac.ir/~fattahi/thesis1.htm>. (Cit. a p. 81).
- Godbold, Natalia. «Beyond information seeking: towards a general model of information behaviour». *Information Research* 11.4. (2006): 269. <<http://InformationR.net/ir/11-4/paper269.html>>. (Cit. a p. 83).
- Guerrini, Mauro e Tiziana Possemato. «Linked data: a new alphabet for the semantic web». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-6305](https://doi.org/10.4403/jlis.it-6305). (2013): 67–90. (Cit. a p. 78).
- Hatop, Götz. «Integrating Linked Data into Discovery». *Code4lib Journal* 21. (2013): 67–90. <<http://journal.code4lib.org/articles/8526>>. (Cit. a p. 85).

- Heath, Tom e Christian Bizer. «Linked Data: Evolving the Web into a Global Data Space». *Synthesis Lectures on the Semantic Web: Theory and Technology* 1.1. (2011): 1–136. <<http://linkeddatatbook.com/editions/1.0>>. (Cit. a p. 78).
- Iacono, Antonella. «Towards a new model of OPAC. From information to knowledge». *JLIS.it* 4.2. DOI: [10.4403/jlis.it-8903](https://doi.org/10.4403/jlis.it-8903). (2013): 85–107. (Cit. alle pp. 79, 81).
- Klein, Maximilian e Kyrios Alex. «VIAFbot and the integration of library data on Wikipedia». *Code4Lib Journal* 2. (2013): 85–107. <<http://journal.code4lib.org/articles/8964>>. (Cit. a p. 86).
- Kroeger, Maximilian. «The road to BIBFRAME: the evolution of the idea of bibliographic transition into a post-MARC Future». *Cataloging & Classification Quarterly* 51.8. (2013): 1–18. (Cit. a p. 80).
- Kulthau, Carol Collier. *Seeking meaning: a process approach to library and information services*. Westport: Libraries Unlimited, 2004. (Cit. a p. 83).
- Le Boeuf, Patrick. «A Strange Model Named FRBRoo». *Cataloging and Classification Quarterly* 50.5-7. (2012): 422–438. (Cit. a p. 81).
- . «Foreword». *Cataloging and Classification Quarterly* 50.5-7. DOI: [10.1080/01639374.2012.682001](https://doi.org/10.1080/01639374.2012.682001). (2012). (Cit. a p. 81).
- Leombroni, Claudio. «Appunti per un'ontologia delle biblioteche digitali: considerazioni sulla Biblioteca digitale italiana». *Bollettino AIB* 44.2. (2004): 115–131. <<http://bollettino.aib.it/article/view/4902>>. (Cit. a p. 97).
- Malmsten, Martin. «Cataloguing in the open - the disintegration and distribution of the record». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-5512](https://doi.org/10.4403/jlis.it-5512). (2013): 417–424. (Cit. alle pp. 79, 85).
- Marchitelli, Andrea e Giovanna Frigimelica. *OPAC*. Roma: AIB, 2012. (Cit. a p. 79).
- Marcum, Deanna. *A Bibliographic Framework for the Digital Age*. 2011. <http://www.loc.gov/marc/transition/pdf/bibframework-10312011.pdf>. (Cit. a p. 80).
- Morando, Federico. «Legal interoperability: making Open Government Data compatible with businesses and communities». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-5461](https://doi.org/10.4403/jlis.it-5461). (2013): 441–452. (Cit. a p. 98).
- Serrai, Alfredo. *Del catalogo alfabetico per soggetti: semantica del rapporto indicale*. Roma: Bulzoni, 1979. (Cit. a p. 89).
- . *Indici, logica e linguaggio: problemi di catalogazione semantica*. Roma: Consiglio Nazionale delle Ricerche - Laboratorio di studi sulla ricerca e sulla documentazione scientifica, 1974. (Cit. a p. 89).
- Žumer, Maja e Marcia Lei Zeng. «ISKO-UK Biennial Conference, July 4-5, 2011, London». 4-5 luglio 2011. <http://www.iskouk.org/conf2011/ppt/zumer.pdf>. (Cit. a p. 92).
- Ventura, Roberto. *Il senso della biblioteca*. Milano: Editrice bibliografica, 2011. (Cit. a p. 89).

- Wenz, Romain. «Linked open data for new library services: the example of data.bnf.fr». *JLIS.it* 4.1. DOI: [10.4403/jlis.it-5509](https://doi.org/10.4403/jlis.it-5509). (2013): 403–415. (Cit. a p. 96).
- Wilson, Thomas D. «Models in information behaviour research». *Journal of Documentation* 55.3. (1999): 249–270. <<http://informationr.net/tdw/publ/papers/1999JDoc.html>>. (Cit. a p. 83).
- Yee, Martha M. «Beyond the Opac: Future Directions for Web-Based Catalogues». *Beyond the OPAC : future directions for Web-based catalogues*. 2006. <http://www.nla.gov.au/lis/stdnrd/grps/acoc/papers2006.html>. (Cit. a p. 80).
- Yu, Holly, Young Margo. «Models in information behaviour research». *The Impact of Web Search Engines on Subject Searching in OPAC* 23.4. (2004): 168–180. <[www.ala.org/ala/lita/litapublications/ital/volume23a/number4a/young.pdf](http://www.ala.org/ala/lita/litapublications/ital/volume23a/number4a/young.pdf)>. (Cit. a p. 83).

ANTONELLA IACONO, Università di Roma "La Sapienza".  
[antonella.iacono@fastwebnet.it](mailto:antonella.iacono@fastwebnet.it)

---

Iacono, A. "Dal record al dato. Linked data e ricerca dell'informazione nell'OPAC". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #9095. DOI: [10.4403/jlis.it-9095](https://doi.org/10.4403/jlis.it-9095). Web.

**ABSTRACT:** The application of linked data to catalogue data opens up new perspectives for OPAC development. The new relationships between bibliographic and other data will produce catalogues that are deeply integrated with the Web, giving a decisive impact on the structure of catalog records, on access and on the functions that the catalog can play in the wider global information space. In this paper the author explores the possibility the new record deconstructed and connected with the other data on the Web is able to facilitate the creation of knowledge in the use of the catalogue. The author then analyzes the potential of application of linked data to the catalogue with regards to the capabilities for research, the new possibilities of semantic search and the ways to access bibliographic data.

**KEYWORDS:** OPAC; Library linked data; Semantic web; Semantic search; Subject search.

---

Submission: 2013-10-03  
Accettazione: 2013-12-10  
Pubblicazione: 2014-01-01





# Links between Libraries and Museums: a Case Study of Library-Museum Collaboration at the Hong Kong Maritime Museum

Patrick Lo, Kitty But, Robert Trio

## 1 Museums versus Libraries. What makes them different from each other?

*Museums, archives and libraries belong at the very heart of people's lives; contributing to their enjoyment and inspiration, cultural values, learning potential, economic prosperity and social equity.*<sup>1</sup>

Both museums and libraries have an important role in cultural preservation and transmission. According to Storey ("Libraries: Their Role and Relationship to Other Cultural Institutions"), "differences among libraries and museums can be profound." The assets, personnel, professional training, and the terminology used can be

---

<sup>1</sup>Impact Evaluation of Museums, Archives and Libraries: Available Evidence Project. Available at: <http://www4.rgu.ac.uk/files/imreport.pdf>.



dramatically different from each other, partly due to the different materials collected and handled by each (Lester). While a 'knowledge artifact' can speak for itself, a 'cultural artifact' requires deeper interpretation. As a result, compared with libraries, museums tend to play a stronger role in providing the context via 'interpretation' (187). Finally, museums have generally a stronger tradition of liaising with the educational community (Allen and Bishoff66). This has led both the museums and libraries to develop distinct techniques and practices for describing, organizing, managing, as well as maintaining their collections (Ray and Choudhury).

However, the museum is also a special type of information system. According to Buckland ("Information as Thing"; "What is a Document?"), objects held in a museum are also considered documents. It is because objects collected and stored inside the museum are also meant to provide information about the object itself, i.e., the type of object, its context and relationships to other objects. From an information-science point of view, museums therefore are very similar to other types of repositories, such as libraries (books and other materials for reading and study) and archives (historical records and documents). However, for a long time, information science has to a large extent neglected the museums. Possibly, it is due to a restricted notion of the nature of information documents to be collected. As explained by Navarrete and Owen,

*books, journals and archival records, even in their digital form, are readily accepted as carriers of information, i.e., as documents. The institutions that provide functions for the collection, storage and access of these documents are generally regarded as information systems in the wider sense, i.e., as the set of human and technical resources, procedures, methods and know-how that together perform one or more specific information functions. ("Museum Libraries: How Digitization*

## Can Enhance the Value of the Museum")

As you can see, although both museums and libraries are considered heritage institutions, and are designed to assemble a single body of knowledge, their distinctive operational practices and standards tend to isolate them from one another. Despite of their similarities and differences, both libraries and museums have a strong educational and recreational role to play, and are both designed to provide resources and a wider range of services to the local community. Recently, there has been greater emphasis on the heritage and leisure function of both museums and libraries (Marsden20), with recognition of a shared common purpose as "cultural heritage institutions working for the public good" (Allen and Bishoff43).

## **2 Relations Between a Museum & Its Museum Library**

Providing the contextual knowledge, and especially general background knowledge, is the function of the museum library. Museum libraries' collections can be made up of books and other archival materials for supporting the research and exhibitions related to the museum subjects. For example, for the museum objects, in order to go beyond the mere admiration and to derive meaningful information from them, one needs the necessary background knowledge (e.g., artistic periods or other technical and cultural information, etc.) and other specific contextual knowledge about the objects (e.g., by whom, when, for what purpose it was created, etc.), and such basic descriptive data is usually provided by the museum curators (Navarrete and Mackenzie Owen14).

On the other hand, the documents held inside the museum library provide general historical, technical, biographical, i.e., infor-

mation pertaining to the collection as a whole, as well as information on individual objects in the collection, and on the connections between individual or collections of objects (Navarrete and Mackenzie Owen). In fact, many museum libraries also collect materials related to the activities held or by the museum, e.g., conservation, restoration, and exhibition, etc. Other materials collected may include books, prints, maps, drawings, music scores, as well as supporting images, e.g., video art, film, postcards, and posters, etc. Occasionally, the museum library collects artists' objects, such as contracts, music scores, manuscripts and personal letters, etc (Navarrete and Mackenzie Owen). As a result, by managing both the museum collections and the museum library, one is able to give access to the "whole" story to the audience, i.e., including both the primary and secondary sources. This greatly enhances the interpretation, understanding as well as enjoyment of the objects and issues for all users, from the very inexperienced novices to the highly knowledgeable specialists.

This specific relationship between the library and the museum objects is the main characteristic that differentiates the museum library from many other types of libraries. Furthermore, for museum libraries, they could be drastically different from each other, in terms of their curators' expertise, origin, collection contents, the way they are organized and funded, and most importantly, the nature and their relations to the parent institutions (i.e., the museum itself). As a result, the format and nature of collaborations between museum curators and librarians need be looked at on a case-by-case. However, it should be highlighted that museum libraries are not always considered part of the resources to be made available to the general public. Taking the Van Abbe Museum Library as an example, it was open to the public by appointment only 20 years after the opening of the Museum. The first librarian was hired 10 years after that (or 30 years after the opening of the museum) so that the Museum Library

could be made open to the public during regular opening hours (Franssen).

### 3 Reasons and Benefits for Ongoing Museum-Library Collaborations

In 2003, library-museum collaboration was described by Diamant-Cohen and Sherman as the “wave of the future” (“Hand in Hand: Museums and Libraries Working Together”). An archivist working in a museum setting, remarked that libraries and archives could learn more from museums about education, while museums could take a page from libraries in providing better access to collections (Wythe).

*It is necessary for collaborators to understand the differences among themselves. A library’s “patron” may be the same individual as a museum’s visitor, but the reason that person has for going to the library may be quite different than the reason for visiting the museum. The library’s funding structure is likely to be very different than the museum’s, and its administrative and decision-making structure may not be at all the same as the museum’s. While the museum has a “catalog” it uses that catalog for inventory purposes, and it may be difficult for a partnering library to understand why that catalog is not available for public use. A library does usually have a preservation program of some kind, but the library may not recognize the tremendous value placed by museums on preservation and conservation handling of collections. Therefore, the library’s primary purpose for starting a digitization project may be to improve access, while the museum will want to include a far greater emphasis on preservation, as might the archival*

*partner. The museum's traditional roles for curators might be compared to the role of university faculty or librarians, but they also are quite different as well. All of us already know these things, but we did not expect them to surface so often in partnership projects between libraries and museums, archives, and historical societies. A new sensitivity to the different systems, cultures, and decision-making requirements is needed.*  
(Allen)

For many years, libraries, museums and archives have collected data in a variety of forms in an attempt to establish the efficiency of service, client use and satisfaction. According to Diamant-Cohen and Sherman (102), Lester (183), Brown and Pollack (212) and Yakel (13), the aims and benefits of library-museum collaboration could be summarized as follows:

- attract new audience groups and expand the reach of the library and museum;
- improve public perceptions of museums and libraries as traditional staid institutions;
- identify new ways to encourage cultural heritage and preservation;
- foster of best practice from the museum and the library;
- share physical resources such as space and materials;
- share policies for preservation and conservation of collections;
- experiencing collaborative working;
- sharing expertise;

- sharing staff training costs.

*By working together, applying collective attention and collaborative action to common challenges and solutions, not only can LAMs leverage network effects, they can also free up valuable time and resources to focus efforts on things only libraries, archives, and museums can do: meet local user's needs, collect unique materials, and preserve those artifacts. (Waibel and Erway)*

Although the benefits of museum-library collaborations are self-evident, non-profit arts and culture groups are challenged as never before. Museums and libraries are without exceptions. Both institutions are constantly competing for audience against a variety of low-cost commercial entertainment products. The operating costs of most non-profit cultural organizations continue to rise. At the same time, we do not see much growth in ticket sales, fund donations and other possible earned income. Most importantly, audiences for the majority of non-profit cultural organizations are rather static, and in many cases, in fact, declining. In short, their audiences have limited attention, and the budgets of most cultural institutions have become increasingly tight. In the face of such of extraordinary challenges and difficulties, how can museums and libraries work together to address the lifelong learning needs of their community, and also working with the communities; and to become the centers for community knowledge creation and development of productive inquiry?

## 4 The Hong Kong Maritime Museum & Its CSSC Maritime Heritage Resource Centre

*Collaboration changes behaviors, processes and organizational structures, and leads to a fundamental inter-connectedness and inter-dependence among the partners, making this transformative change the hallmark sign by which true collaboration can be known. (Waibel and Erway)*

*Museums and libraries can continue to enrich the education and knowledge building of society by consciously adapting their services to the new ways in which people want to seek, gather, and interpret information and to create meaning in the 21<sup>st</sup> century. (Pastore)*

Established in 2005, the Hong Kong Maritime Museum (HKMM)<sup>2</sup> is a relatively new museum that acquires, conserves, researches and exhibits all forms of materials related to the development of boats, ships, maritime exploration and trade, and naval warfare on the South China coast as well as its adjacent seas. This Museum aims to provide a comprehensive account of Hong Kong's growth and developments as a major world port and a maritime centre. The HKMM includes many semi-permanent and special exhibitions (e.g., the Canton Trade with Europe and the Americas,<sup>3</sup> Piracy in the China Seas,<sup>4</sup> etc.), interactive displays, educational events, café, and

---

<sup>2</sup><http://www.hkmaritimemuseum.org/>.

<sup>3</sup><http://www.hkmaritimemuseum.org/eng/explore/galleries/the-canton-trade-with-europe-and-the-americas/12/6/>.

<sup>4</sup><http://www.hkmaritimemuseum.org/eng/explore/galleries/piracy-in-the-china-seas/12/12/>.

also a museum shop.<sup>5</sup> The HKMM does not only serve as a unique platform for teaching and learning of the local heritage in Hong Kong, it also contributes greatly to the promotion of community engagement and social connections. The HKMM is also equipped with its own museum library, and it is called the CSSC (China State Shipbuilding Corporation<sup>6</sup>) Maritime Heritage Resource Centre.<sup>7</sup> In addition to supporting various research activities carried out by the Museum, this Resource Centre also serves as a central and yet comprehensive repository for different publications and archival documents on maritime heritage and history related to Hong Kong, as well as Southeastern China. Some of these items actually date back more than 300 years, and represent work that cannot be found in the public domain anywhere else in Hong Kong.<sup>8</sup>

*Participants identified museums and libraries as purpose-driven institutions - socially responsible and socially responsive. Museums and libraries already act as safe communal spaces for people to interact with one another. As opportunities for social engagement outside of private or working life and removed from the profit interests of commercial spaces, museums and libraries have the ability to identify and respond to community needs in ways that other spaces cannot. (9)*

*Collaboration is essential to the sustainability and future growth of museums and libraries and can increase the ability*

---

<sup>5</sup><http://www.hkmaritimemuseum.org/eng/explore/galleries/introduction/12/4/>.

<sup>6</sup>The Resource Centre is named after China State Shipbuilding Corporation in recognition for their support of the Centre's staff and resources.

<sup>7</sup><http://www.hkmaritimemuseum.org/eng/resource-centre/introduction/about-us/30/59/>.

<sup>8</sup><http://www.hkmaritimemuseum.org/eng/resource-centre/introduction/about-us/30/59/>.

*of institutions to address 21<sup>st</sup> century challenges and opportunities and meet community needs. (17)*

According to Kitty But and Robert Trio, the HKMM and the Resource Centre on both sides are actively gaining new facilities, new buildings, developing new collections, launching new education programmes, and improving access to collections through digitization. A large amount of resources are being invested in promoting the educational and heritage roles of the Museum and the Resource Centre museums. In short, collaboration has made it possible to set up many new education initiatives and outreach programmes, which would have been made impossible without the museum-library partnership (see interview transcript in Appendix for examples and other details).

In spite of all these new and drastic changes, the unique and yet core responsibilities of both institutions (the library and the museum) will remain the same, i.e., to continue to acquire and to preserve the original materials as well as other relevant information, and to make them as widely accessible as possible.

To summarize the key issues addressed in the interview, the aims, objectives, services and operations of both libraries and museums do and can fit well together, especially those concerned with lifelong learning, heritage preservation and social inclusion, etc. As pointed out by Pastore (*The Future of Museums and Libraries: A Discussion Guide. (IMLS-2009-RES-02)*),

*while museums and libraries as distinct entities may have equally distinct approaches to dealing with 21<sup>st</sup>-century issues, they also have common ground to stand on as organizations in the public trust and for the public good.*

If museum curators and librarians begin to collaborate more in the future, they may also begin to re-define themselves, i.e., their new

educational and recreational roles in the local context and in the global knowledge society. As indicated in the interview above, there is no single concept of what a museum library should be, and there is no single solution to all problems currently faced by museum curators and librarians. Supportive to the ideology pointed out by Gibson, Morris and Cleeve (59), museums and libraries need to work together “to create a brand new model for learning in an information-rich environment by taking the best aspects of two informal learning environments and blending them to create a seamless integration of resources.”

The new emphasis on museum-library collaborations will be carried across curation, education, marketing and even fundraising, and thereby replacing many traditional divisions of museum and library practices. Both museums and libraries are in a unique position to provide a rich experience for the visitors, and to engage the public in terms of their rich and exciting contents. The new collaborative efforts aim at providing a much boarder and experience that seek to be part of the lifestyle of its audiences, and a trusted source for their cultural and intellectual growth for the community as a whole. In order to achieve this, museums need to change their thinking about how to integrate both museum and library services in a digital context, thereby enhancing the informational value of the museum in important ways.

Perhaps, the greatest concern for museums and libraries today is - what they need to do to remain meaningful institutions in the future. In other words, how museums and libraries could stay relevant to the communities in which they reside and serve; and how to best preserve and provide access to the materials in a way that serves future generations. Such intimate partnership strategies via digitization does not only help engage the institutions in both physical and virtual communities, they also further promote the

preservation of the institutions' unique identity and resources - as an integrated living and breathing institution that actively publicize its role in the key contemporary issues, such as cultural heritage preservation. Since information contents and services over the Internet are increasingly syndicated for increased interoperability and connectivity, curators and librarians may take full advantage of this by identifying ways to enable the cultural collections to leverage the power of Google-, Amazon-, flickr- and Facebook-dominating networked environment.

*In The Great Good Place, Ray Oldenburg defines the significance of the third place in a healthy society. Neither work nor home, the third place is a neutral community space, where people come together voluntarily and informally in ways that level social inequities and promote community engagement and social connection. As public gathering places organized around public service and the transfer of information and ideas across individuals, museums and libraries are a unique form of the third place because of their distinct resources as easily accessible, low-cost barrier places rich in content and experience.*  
(Pastore)

## 5 Conclusion

To conclude this paper, we would like to reiterate the importance of having both museums and libraries working closely together - to create engaging, pleasant and memorable experiences for their visitors, so that they could be fully inspired, enriched as well as entertained in their search for knowledge. Michael Fox (Minnesota Historical Society) expressed a sentiment during the RLG Forum,

*I continue to argue that good museums need to become more*

*like research libraries and archives just as good libraries and archives ought to adapt certain characteristics of the museum experience.*

While “learning from each other” was presented as an opportunity, “lending to each other” was established as a real-life example of museum and library working hand-in-hand. As collaboration, collocation, and the blending of services continue, it is very likely that the boundary between museum and library will become increasingly less distinct in the future. However, this does not mean that the unique mission and identity of an individual institution must disappear; rather, the institutional goals can be expanded, and enhanced through ongoing collaborations that combine resources, knowledge, and experience for mutual benefits (9).

## Postscript

### Collecting conflicts

Libraries and museums have more in common than they are different. But it is the small differences that often lead to conflicts from an operational point of view. These conflicts arise especially when an organization has a shared space with a library and a museum. Although the library at the Hong Kong Maritime Museum acts as an independent office of the museum and follows standards and best practices of a library system, the library is heavily influenced by the mission and vision of the museum and its core practices.

One can see from the HKMM library’s current collection that it was predisposed in the early years not to collect not in a typical library fashion. The library’s primary focus on collecting had been for first editions and rare copies. There was never an intention to lend items or have a juvenile section. The collection grew primarily

from the wishes of one person who approached the library from a researcher's point of view. And despite this focus, there were few reference materials collected. Collecting rare and first editions exclusively is consistent with how museum professionals think about collecting artifacts for the collection.

At HKMM this practice has led to some interesting dynamics when it comes to cataloging collections. Libraries tend to focus on cataloging records based on the idea that the item is identical to similar items of the same edition or printing. Museums tend to catalog items as if it is a one of a kind. And the item's individual story may be tied to a person who owned it or from it was created. So when a dual institution acquires a collection, difference of opinions often arise on how the item should be cataloged and what type of metadata should be collected on it.

Because personal histories are an important part of museums' missions, it is this concept that often drives the thinking behind how cataloging is done. Museums cataloging places the connection of the person to the object above all other things. It is the personal connection that provides the context of why the object is in the collection. For example, a printed copy of a Judaic prayer book from 1930 takes on a much greater meaning if that book belonged to a Holocaust survivor.

### **Case example**

Recently a small group of nautical charts came into the collection of the museum. In many ways, this set of charts helps to illustrate the struggle that exists in a museum setting that also a library focus. The charts also illustrate the struggle that institutions that tell personal stories have when the materials are not yet old enough to be considered antiques but may be unique enough that in the future one may want to research because they belonged to a specific person.

The British Hydrographic Office produced this set of charts in the 1970's. Sailors of small vessels commonly used these charts for short hops from one island to another. The charts show landmasses and the depth of the water in stunning detail. In all, there were about 40 charts covering most of the major waterways of Hong Kong. This practice of using charts has now been replaced by modern GPS systems that are on board most small vessels.

The rub that soon ensued amongst museum staff was where to place this collection: the library, the museum or the archive. The librarian made a case that the charts were clearly a secondary source and fit under well-established library cataloging practices. Some argued that the charts all belonging to this one sea captain belong in the archives. In the future one may wish to research this particular person and the charts represent his life work. Lastly, because the charts contained unique notations and markings from this sea captain, they could be considered a museum collection item. The sea captain had also donated a set of nautical instruments to the museum. By making the charts as a part of the museum collection, all of these items could be stored together.

## Resolution

The challenge for the museum was to preserve the set together. In the end, it was the concept that the personal story was the most important. The museum may never know if a researcher may wish to explore this person's life work but institutions like the Hong Kong Maritime Museum must ensure that it is possible if they do.

## References

Allen, Nancy. "Collaboration Through the Colorado Digitization Project". *First Monday* 5.6. (2000). (Cit. on p. 108). Web.

- Allen, Nancy and Liz Bishoff. "Collaborative Digitization: Libraries and Museums Working Together". *Advances in Librarianship* 26. (2002). (Cit. on pp. 104, 105). Print.
- Brown, Karen and Miriam Pollack. "Illinois Libraries and Museums: Connecting and Collaborating for the Future". *Illinois Libraries* 82.3. (2000). (Cit. on p. 108). Print.
- Buckland, Michael K. "Information as Thing". *Journal of the American Society of Information Science* 42.5. (1991): 351–360. (Cit. on p. 104). Print.
- . "What is a Document?". *Journal of the American Society of Information Science* 48.7. (1997): 804–809. (Cit. on p. 104). Print.
- Diamant-Cohen, Betsy and Dina Sherman. "Hand in Hand: Museums and Libraries Working Together". *Public Libraries* 42.2. (2003): 102–105. (Cit. on pp. 107, 108). Print.
- Franssen, Fiana. "Van boekenplank naar kennisproductie: De geschiedenis van de bibliotheek van het Van Abbemuseum 1936-2007". *Kunstabibliotheken in Nederland*. Ed. Roman Koot, Michiel Nijhoff, and Saskia Scheltjens. Leiden: Primavera Pers, 2007. 99–112. (Cit. on p. 107). Print.
- Gibson, Hannah, Anne Morris, and Marigold Cleeve. "Links between Libraries and Museums: Investigating Museum-Library Collaboration in England and the USA". *Libri* 57. (2007). (Cit. on p. 113). Print.
- Lester, Ray. "Convergence of Museums and Libraries?". *Alexandria* 13.3. (2001): 183–191. (Cit. on pp. 104, 108). Print.
- Marsden, Christopher. "Sectors and Domains: Some Reflections on Co-Operation and Integration". *Journal of the Society of Archivists* 22.1. (2001): 17–23. (Cit. on p. 105). Print.
- Navarrete, Trilce and John Mackenzie Owen. "Museum Libraries: How Digitization Can Enhance the Value of the Museum". *Palabra Clave (La Plata)* 1.1. (2011). (Cit. on pp. 104–106). Web.
- Pastore, Erica. *The Future of Museums and Libraries: A Discussion Guide*. (IMLS-2009-RES-02). Washington, D.C.: Institute of Museum and Library Services, 2009. (Cit. on pp. 110–112, 114, 115). Web.
- Ray, Joyce and Sayeed Choudhury. "Web-Wise 2002 - Conference on libraries and museums in the digital world: a report by the co-organizers". *Library Hi Tech News* 19. (6 2002). (Cit. on p. 104). Web.
- Storey, Tom. "Libraries: Their Role and Relationship to Other Cultural Institutions". *OCLC Newsletter* 260. (April 2003): 12–14. (Cit. on p. 103). Print.
- Waibel, Günter and Ricky Erway. "Think Global, Act Local - Library, Archive and Museum Collaboration". *Museum Management and Curatorship* 24.3. (2009). (Cit. on pp. 109, 110). Web.

Wythe, Deborah. "New Technologies and the Convergence of Libraries, Archives, and Museums". *RBM: a Journal of Rare Books, Manuscripts and Cultural Heritage* 8.1. (2007). (Cit. on p. 107). Print.

Yakel, Elizabeth. "Choices and Challenges: Cross-Cutting Themes in Archives and Museums". *OCLC Systems & Services: International Digital Library Perspectives* 21.1. (2005): 13–17. (Cit. on p. 108). Web.

PATRICK LO, University of Tsukuba.  
[wotan455@hotmail.com](mailto:wotan455@hotmail.com)

KITTY BUT, The Hong Kong Maritime Museum.  
[kittybut@hkmaritimemuseum.org](mailto:kittybut@hkmaritimemuseum.org)

ROBERT TRIO, The Hong Kong Maritime Museum.  
[rstriojr@gmail.com](mailto:rstriojr@gmail.com)

---

Lo, P., K. But. R., R. Trio. "Links between Libraries and Museums: a Case Study of Library-Museum Collaboration at the Hong Kong Maritime Museum". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #8916. DOI: [10.4403/jlis.it-8916](https://doi.org/10.4403/jlis.it-8916). Web.

ABSTRACT: Established in 2005, the Hong Kong Maritime Museum (HKMM) is a relatively new private museum that aims to collect all forms of materials related to the development of boats, ships, maritime exploration and trade, and naval warfare on the South China coast, as well as its adjacent seas. The Museum not only serves as a unique platform for teaching and learning of the local heritage in Hong Kong, it also contributes greatly to the promotion of community engagement and social connections. The HKMM is also equipped with its own museum library, and it is called the CSSC (China State Shipbuilding Corporation) Maritime Heritage Resource Centre. In addition to supporting various research activities carried out by the Museum, this Resource Centre also serves as a central, and yet comprehensive repository for publications, and other archival documents on maritime heritage and history related to Southeastern China.

This paper aims to compare the distinctive operational practices, and user needs between museums and libraries. It also examines the benefits and challenges of museum-library collaborations in the new knowledge-driven society. This paper

features an interview with Kitty But (Librarian, CSSC Maritime Heritage Resource Centre, The Hong Kong Maritime Museum) and Robert Trio (Project Officer for Technology, The Hong Kong Maritime Museum), and in which they discussed their professional experiences in the fields of audience education; the implementation of different new technologies associated with the museum and library services; and various collaborative initiatives carried out between the Museum and the Resource Centre. Upcoming challenges and opportunities faced by both the Museum and Resource Centre are also discussed in this paper..

KEYWORDS: HKMM; Hong Kong Maritime Museum; Library-Museum Collaboration; Operational practices; User needs.

---

Submitted: 2013-05-23

Accepted: 2013-07-05

Published: 2014-01-01





# Preservare la letteratura scientifica elettronica in open access

Maria Teresa Biagetti

La preservazione dei documenti digitali prodotti dall'editoria elettronica commerciale, volta alla salvaguardia dell'uso a lungo termine dei documenti che potrebbero essere non più disponibili, anche a causa di situazioni indipendenti dalla volontà degli editori, è una problematica ampiamente affrontata e continuamente riproposta all'attenzione dei bibliotecari e degli analisti delle procedure per la conservazione. In particolare, la diffusione delle riviste scientifiche *e-only*, edite nel solo formato elettronico dagli editori, e sottoposte al regime delle licenze di utilizzazione, che costringe le biblioteche ad usare i contenuti gestiti in sistemi controllati dagli editori, ha reso urgente l'adozione di misure di preservazione della letteratura scientifica, soprattutto per le biblioteche accademiche, per assicurare nel futuro la fruizione dei risultati delle ricerche. Il potenziale rischio costituito dal legame con la gestione editoriale dei contenuti, fu evidenziato con vigore in un documento sottoscritto il 13 settembre 2005 da diverse biblioteche di università statunitensi, da MIT e da New York Public Library, intitolato *Urgent Action Needed to Preserve Scholarly Electronic Journals* (Waters), nel quale si tracciavano le linee programmatiche della politica per l'archiviazione condivisa delle riviste scientifiche da parte delle biblioteche accademiche in archivi qualificati. Veniva posto infatti l'accento proprio sulla necessità che



le biblioteche delle università riconoscessero la preservazione delle riviste elettroniche in depositi certificati come una forma di assicurazione, e una modalità di gestione di due distinte tipologie di rischio di perdita dei contenuti: da un lato il rischio della perdita totale e permanente delle riviste stesse, dall'altro lato il rischio per le biblioteche di subire l'interruzione della fruizione, per periodi più o meno lunghi, a causa del fallimento di un editore, rischi da scongiurare allestendo archivi certificati fuori del controllo editoriale.<sup>1</sup>

D'altra parte, anche i documenti scientifici pubblicati nell'ambito della filosofia dell'open access<sup>2</sup> rischiano di non essere disponibili, ad esempio nel caso in cui si verifichino danni alla piattaforma di distribuzione, o nel caso di eventi catastrofici, ma anche di banali errori e disfunzioni durante il passaggio ad un diverso istituto gestore, eventi che possono provocare l'impossibilità di accedere ai documenti, seppure temporaneamente.<sup>3</sup> Anche su di essi, quindi, è necessario porre l'accento, affinché si determini una maggiore consapevolezza della necessità della preservazione in molteplici depositi di garanzia e si adottino misure di sicurezza ridondanti, dal momento che, in questi casi, la pubblicazione provvidenzialmente realizzata anche nel formato cartaceo, avviene in genere con una tiratura molto limitata, contando proprio sulla disponibilità dell'accesso aperto, e di conseguenza la diffusione non è capillare.

---

<sup>1</sup>«First, research and academic libraries and associated academic institutions must recognize that *preservation of electronic journals is a kind of insurance*, and is not in and of itself a form of access. Preservation is a way of managing risk, first, against the permanent loss of electronic journals and, second, against having journal access disrupted for a protracted period following a publisher failure» (Waters).

<sup>2</sup>La letteratura sull'open access è vastissima, ed è impossibile citarla qui (per una bibliografia aggiornata, rinvio al lavoro curato da Bailey).

<sup>3</sup>Sono stata spinta ad approfondire questa problematica avendo subito la sparizione dalla consultazione in open access, per più di un anno, di un mio articolo – e del fascicolo sul quale era stato pubblicato – nel momento in cui la rivista ha utilizzato una diversa piattaforma per l'accesso aperto.

È importante ricordare che la *Berlin Declaration on Open Access to knowledge in the Sciences and Humanities*,<sup>4</sup> la dichiarazione firmata a Berlino il 22 ottobre 2003 da vari paesi, tra cui Germania, Italia, Francia, Norvegia, Ungheria, cita come imprescindibile l'archiviazione in un formato standard e la preservazione a lungo termine dei contributi offerti ad accesso aperto via Internet.<sup>5</sup> Una copia del lavoro, nella versione considerata definitiva dall'autore, insieme all'autorizzazione a riprodurlo, utilizzarlo, distribuirlo, deve essere infatti depositata almeno in un archivio consultabile in rete, che sia mantenuto da una istituzione, scientifica o dipendente da una agenzia governativa, o comunque una associazione affidabile, che ne assicuri sia la distribuzione senza limiti, sia l'archiviazione a lungo termine.

Prima di entrare nel merito della preservazione in archivi qualificati della letteratura scientifica pubblicata seguendo la strategia dell'accesso aperto, ed esaminare alcune realizzazioni che offrono anche questa possibilità alle biblioteche, può essere interessante soffermarsi ad analizzare alcuni aspetti meno conosciuti del fenomeno stesso dell'open access, dalla sua non massiccia diffusione, alla modesta utilizzazione dei repository istituzionali da parte dei ricercatori, anche nel settore della Library and Information Science e, d'altra parte, all'alta incidenza, invece, dell'utilizzazione della letteratura scientifica edita attraverso questa modalità.

---

<sup>4</sup><http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung>. Su «JLIS.it», 3.2 (2012), <http://leo.cilea.it/index.php/jlis/issue/view/542>, sono ora disponibili tutti i documenti, alcuni anche in traduzione italiana.

<sup>5</sup>Per la presentazione dettagliata delle dichiarazioni sottoscritte nell'ambito dell'open access, v. Guerrini; Cassella, *Open access e comunicazione scientifica. Verso un nuovo modello di disseminazione della conoscenza*.

## **1 Livelli di diffusione e tasso di citazione della letteratura scientifica in open access**

Le riviste scientifiche che hanno adottato la modalità della libera diffusione in rete, con il controllo della qualità, e si sono inserite successivamente nel più complesso movimento dell'open access, hanno raggiunto la massima diffusione nei primi anni di questo secolo negli USA, in Canada, in Europa, e in un secondo momento in India e in Australia, in parallelo con il sorgere di repository istituzionali o su base disciplinare che, a cominciare dagli anni Novanta, hanno mostrato una rapida espansione concentrandosi, nell'ultimo decennio trascorso, soprattutto negli USA, in Europa, in India, in Giappone e in Cina (Xia).

Il fattore che ha contribuito maggiormente a stimolare lo sviluppo sia delle riviste che dei repository ad accesso aperto, viene generalmente individuato nella maggiore diffusione dell'utilizzazione della rete, che anche nei paesi in via di sviluppo risulta consistente almeno nei centri urbani, dove si concentrano Università e centri di ricerca. Tra il 2000 e il 2010 l'uso del web si è progressivamente allargato, dagli USA e Canada, verso Australia ed Europa, e poi verso Sud Africa e Russia.

La diffusione dell'utilizzazione della rete e l'impiego della tecnologia informatica non sono comunque l'unico fattore che spiega lo sviluppo della filosofia dell'accesso aperto: ad esso si aggiungono le caratteristiche socio-culturali e politiche dei singoli paesi. Il Giappone, ad esempio, rispetto ai vicini Cina e Corea del Sud, ai quali è paragonabile dal punto di vista tecnologico, presenta un numero considerevolmente più alto di repository istituzionali. La spiegazione risiede in gran parte nella tradizionale consuetudine degli studiosi giapponesi, a differenza di quelli cinesi, e molto prima del sorgere del movimento dell'open access, di scambiarsi le pubbli-

cazioni attraverso riviste scientifiche, denominate Kyo, liberamente distribuite da Università e istituzioni, sviluppatesi prima nell'ambito delle scienze umane – e questo è singolare – e poi diffuse tra le scienze esatte: questa consuetudine ha predisposto i giapponesi ad accogliere l'accesso aperto molto più facilmente, quando questo modello si è imposto (*Ivi.*, p. 90).

Uno studio realizzato da Thomson Reuters Corp. nel 2004 (McVeigh), con l'obiettivo di analizzare quantitativamente la crescita e la diffusione delle riviste in open access nei diversi settori disciplinari, ha considerato la presenza delle riviste in open access all'interno dei database citazionali gestiti da ISI Thomson, e l'impatto dell'OA sulle citazioni, rilevando come appena il 2.6% delle riviste presenti in Web of Science fossero ad accesso aperto, la stragrande maggioranza appartenessero ai settori della Medicina e della Biologia, e le Scienze umane e sociali fossero modestamente rappresentate. Oltre a ciò, è stato però messo in evidenza che più del 55% delle riviste trattate in ISI WoS nel 2003, sono riconducibili a editori che permettono una qualche forma di auto-archiviazione, e fanno cioè parte del circuito che accoglie le indicazioni del movimento per l'open access, e quindi in quelle riviste sono pubblicati articoli potenzialmente auto-archiviabili dagli autori.

Un'altra indagine, con cui sono stati censiti gli articoli editi in accesso aperto nel 2008 in tutte le discipline coperte da Scopus, reperiti tramite Google, eliminando le copie occasionalmente presenti in rete e accettando invece i pre-print, ha fornito come risultato una disponibilità media degli articoli in open access complessivamente del 20,4%, e in particolare, del 8,5% per quanto riguarda le riviste scientifiche e 11,9% per i repository istituzionali e i siti, distribuiti in maggioranza tra discipline quali Life sciences, Earth sciences, Biochemistry ecc, con la sola presenza delle Social sciences al di fuori delle discipline "dure", e la totale mancanza delle Scienze umane (Björk).

Nonostante l'indubbia espansione della nuova modalità di pubblicazione anche nei repository appositamente creati e mantenuti attivi, la *green road*, alcune indagini sull'uso che le comunità accademiche hanno fatto nell'ultimo decennio dei loro depositi istituzionali, nella cui gestione sono molto spesso coinvolte proprio le biblioteche degli Atenei, hanno consentito di evidenziare che la pubblicazione di pre print o post print nei repository non costituisce sempre il canale principale e preferito della comunicazione scientifica, ma che invece permane un legame forte con il modello tradizionale di divulgazione dei risultati della ricerca, nonostante gli indubbi vantaggi offerti dal sistema dell'accesso aperto: disseminazione immediata, maggiore utilizzazione e diffusione, più alto livello di citazioni (Cullen e Chawner). Anche nel settore della *Library and Information Science*, sorprendentemente, l'auto-archiviazione non risulta essere una pratica regolare. Da un'indagine su un campione di 922 articoli, pubblicati su 20 riviste del settore, ad alto fattore d'impatto, e la cui disponibilità in accesso aperto è stata verificata attraverso Google Scholar, è risultata una percentuale di appena il 27% (Way). Inoltre, dall'indagine è emerso che anche i ricercatori di questo settore disciplinare depositano con riluttanza i propri articoli nei repository istituzionali e disciplinari, e in ogni caso si limitano ad archiviare le ricerche in un solo deposito, dimostrando scarsa consapevolezza dell'importanza della ridondanza delle localizzazioni al fine di una sicura preservazione.

Un elenco di repository open access di Università e Enti di ricerca, la cui qualità è stata controllata, suddivisi per continenti e per nazioni, è offerto in *OpenDOAR, The Directory of open access repositories*,<sup>6</sup> un servizio realizzato da SHERPA, il consorzio cui partecipano molte Università e istituti di ricerca inglesi, oltre alla British Library, coordinato dall'Università di Nottingham. Offre la possibilità di

---

<sup>6</sup><http://www.opendoar.org>.

eseguire ricerche su tutti i depositi, oppure solo su quelli istituzionali, o disciplinari o governativi, o aggregati, e di compiere ricerche per tipologia di contenuti. A luglio 2013 risultano per l'Italia 73 archivi aperti, 65 dei quali istituzionali, appartenenti ad Università e Enti di ricerca. Indagini condotte utilizzando come fonte proprio OpenDOAR, hanno messo in luce però che, ad esempio nel Regno Unito, la maggior parte dei depositi istituzionali di tipo accademico gestisce meno di 10.000 documenti, solo in una modesta percentuale in full-text, la cui qualità non sempre viene garantita da *referee*. Le forme di auto-deposito da parte degli studiosi sono maggiormente diffuse in quei campi disciplinari di recente costituzione, o che necessitano di una aggregazione rapida delle forze scientifiche per imporsi sullo scenario internazionale, mentre lo sono in misura minore nei settori disciplinari più solidi, nei quali gli studiosi possono contare su modelli di comunicazione radicati, su riviste scientifiche dal prestigio riconosciuto e su circuiti di disseminazione efficaci (Cullen e Chawner). Nei casi in cui i ricercatori adottino l'auto-archiviazione, ai repository disciplinari (*subject based*) vengono in genere preferiti i depositi istituzionali, gestiti dalle Università o da istituzioni governative di ricerca. In entrambi i casi frequentemente vengono usati software open source come E-Prints sviluppato dall'Università di Southampton, per gestire l'archiviazione di libri, articoli, working papers, tesi, interventi a conferenze, ecc., compatibile con Open Archive Initiative (OAI) e predisposto per l'*harvesting* dei metadati, o anche DSpace (prodotto da MIT), più flessibile di E-Prints, in quanto accoglie set di dati, files video e audio, programmi per computer, e permette anche la preservazione a lungo termine. In definitiva, si può affermare che appare comunque piuttosto lontana dalla realtà, l'idea che i depositi istituzionali della produzione scientifica possano oggi costituire l'unico canale privilegiato per la diffusione della ricerca scientifica.

D'altra parte, se consideriamo invece l'impatto prodotto sulla ricerca scientifica dalla letteratura presente liberamente online, determinabile attraverso il computo delle citazioni ricevute sia dagli articoli pubblicati sulle riviste che hanno adottato la strategia dell'open access, sia dalla produzione resa accessibile attraverso il deposito in repository disciplinari o istituzionali, possiamo rilevare che questo è generalmente molto più alto di quello determinato dagli studi pubblicati su riviste in formato cartaceo, oppure in formato elettronico, ma non liberamente accessibili.<sup>7</sup>

Kristin Antelman ha preso in considerazione i dati relativi alle citazioni ricevute da un campione di articoli disponibili ad accesso libero nelle aree della Matematica, delle Scienze politiche, dell'Ingegneria elettrotecnica ed elettronica e della Filosofia, e li ha comparati poi con quelli relativi agli articoli non in open access. Le citazioni sono state desunte da ISI Web of Science, eliminando le auto-citazioni, e sono state scelte le riviste con il più alto fattore d'impatto, a sua volta basato sul numero delle citazioni ricevute. La ricerca degli articoli liberamente disponibili è stata poi ampliata attraverso il motore di ricerca Google, recuperando un'altra importante quota di full-text in libera consultazione come drafts, pre-print o post-print, depositati sulle pagine personali degli autori e sui siti istituzionali, e adottando in effetti una concezione di open access piuttosto elastica ed ampia. Dalla ricerca emerge che, in particolare per il settore della Matematica, gli articoli pubblicati in open access, o comunque disponibili liberamente sui siti degli autori, sono enormemente più citati di quelli non *free* e quindi hanno un maggiore impatto globale sulla ricerca: in sostanza, si verifica un aumento delle citazioni che va dal 45% in più nel campo della Filosofia, al 91% in più in quello della Matematica. Il numero totale degli articoli accessibili

---

<sup>7</sup>Alcune considerazioni su questa tematica, anche nell'articolo di De Robbio, «Analisi citazionale e indicatori bibliometrici nel modello Open Access».

liberamente, in realtà, è molto probabilmente sottostimato: infatti bisogna considerare gli articoli *ephemeral*, temporaneamente sicuramente disponibili online, ma non presenti in rete al momento della ricerca perché rimossi per qualche motivo, ma anche la possibilità che alcuni articoli, offerti alla libera consultazione come pre-print, siano successivamente stati rimossi dall'autore o sostituiti con il link all'articolo, pubblicato in modalità non *free* dall'editore commerciale. A questo bisogna aggiungere il fatto che alcuni articoli potrebbero avere subito leggere modifiche nei titoli nel passaggio dal pre-print alla pubblicazione definitiva, e quindi non siano stati recuperati solo per questo motivo («Do Open-Access Articles Have a Greater Research Impact?»). Questi dati sono stati sostanzialmente confermati da quelli prodotti da una vasta indagine, che ha riguardato ben dieci discipline, tra cui Economia, Psicologia, Scienze politiche, Biologia e Giurisprudenza, e gli articoli editi sulle riviste peer-reviewed pubblicate tra il 1992 e il 2003, desumendo i dati delle citazioni da ISI. Nei 12 anni di riferimento, la percentuale di articoli in open access nelle discipline considerate varia, a seconda delle discipline, dal 5% al 16%, e l'aumento delle citazioni ricevute dagli articoli considerati varia, sempre relativamente alle discipline, dal 36% al 172% (Hajjem, Harnad e Gringras).

L'aumento del tasso delle citazioni potrebbe essere spiegato con l'attività di auto-selezione da parte degli autori, i quali tenderebbero ad auto-archiviare i lavori migliori, che quindi saranno certamente più citati. Tuttavia, questa spiegazione non dovrebbe essere valida nei settori disciplinari in cui l'auto-archiviazione raggiunge quote del 60% o 70%, come in Astrofisica, e invece anche in queste discipline l'aumento delle citazioni rientra nel medesimo *range* (Hajjem, Harnad e Gringras). L'esame comparativo dell'auto-archiviazione decisa dai singoli autori (*self-selective self-archiving*), i quali stabiliscono quale tra i propri articoli destinare alla libera consultazione

online, e dell'auto-archiviazione imposta come obbligo dagli enti di ricerca o dalle Università (*mandatory self-archiving*), ha permesso di verificare che l'aumento delle citazioni si determina sia nell'uno che nell'altro caso, vanificando la possibilità di attribuire esclusivamente alle scelte degli autori, che pongono in accesso aperto i loro lavori migliori, il motivo dell'incremento delle citazioni (Gargouri et al.).<sup>8</sup>

Un'altra ricerca ad ampio raggio, dedicata ai campi disciplinari della Matematica applicata, dell'Economia, dell'Ecologia e della Sociologia, che ha adottato come strumenti per verificare la disponibilità ad accesso aperto sia Google Scholar che Google e, meno fruttuosamente, OAster e OpenDOAR, ha permesso di rilevare come il 49% dei 4.633 articoli che componevano il campione fossero ad accesso aperto, e che il loro tasso medio di citazione raggiungeva il 9.04, contro il 5.76 degli articoli reperibili a pagamento, considerando nel computo delle citazioni anche le auto-citazioni, effettuate sia sulla medesima rivista su cui era stato pubblicato l'articolo citato, sia su una rivista diversa, risultate comunque in numero nettamente inferiore rispetto alle citazioni ricevute da autori completamente estranei (Norris, Oppenheim e Rowland). Sebbene la diffusione in ambiente scientifico delle modalità di pubblicazione ad accesso aperto risulti, dalle indagini qui considerate, non capillare, è invece proprio l'alto tasso di utilizzazione degli articoli che presentano i risultati delle ricerche scientifiche, come si evince dalle analisi delle citazioni ricevute, a convincere della necessità di azioni più decise, che contemplino la preservazione a lungo termine in una pluralità di depositi certificati, della letteratura pubblicata in open access.

---

<sup>8</sup>In questo caso, è stato analizzato un campione di 27.197 articoli, scelti tra quelli pubblicati tra il 2002 e il 2006 su 1.984 riviste, e composto da articoli in open access per scelta dei loro autori (generalmente sui propri siti personali) e articoli obbligatoriamente in open access per decisione istituzionale.

## 2 Preservazione dei documenti scientifici ad accesso aperto e *dark archives*

Le riviste scientifiche elettroniche prodotte dall'editoria commerciale, oltre ad essere dipendenti dalle tecnologie, che mutano ed evolvono continuamente, come è ampiamente noto, sono soggette ad un'accessibilità e ad un'utilizzazione incerte e limitate ai periodi stabiliti dagli accordi editoriali. È necessario quindi che le biblioteche archivino e preservino le annate delle riviste di rilievo scientifico, in modo che siano sicuramente accessibili nel futuro, e possano essere a disposizione dei ricercatori.<sup>9</sup>

In alcuni casi sono le biblioteche nazionali ad assumersi il compito di preservare le riviste elettroniche, collegandolo alla funzione di centro per il deposito legale dei documenti elettronici. In Italia, ad esempio, il progetto Magazzini Digitali,<sup>10</sup> coordinato dalla Fondazione Rinascimento Digitale,<sup>11</sup> in collaborazione con la Biblioteca nazionale centrale di Firenze, la Biblioteca nazionale centrale di Roma e la Biblioteca nazionale Marciana, e diretto dal 2006 da Giovanni Bergamin, ha realizzato un'infrastruttura informatica persistente per gestire e preservare a lungo termine i documenti elettronici pubblicati in Italia e diffusi tramite rete informatica. I dati vengono replicati sui tre diversi siti corrispondenti alle tre biblioteche nazionali coinvolte, e tra queste, la nazionale Marciana costituisce il *dark archive*, che conserva due copie di emergenza di ciascun file, su due macchine diverse. Dal 2010 il MiBac sperimenta il deposito legale nazionale dei documenti digitali nativi: e-journals pubblicati da editori commerciali, e Tesi di Dottorato (Bergamin e Messina). La biblioteca

---

<sup>9</sup>Il problema è stato ampiamente affrontato anche nella letteratura professionale italiana. Un primo esame in Giordano; maggiori dettagli in Kirchoff e Gifford Fenton; Cassella, «Le biblioteche nella transizione al "solo elettronico"».

<sup>10</sup><http://rinascimento-digitale.it/magazzinidigitali.phtml>

<sup>11</sup><http://rinascimento-digitale.it/home.phtml>

nazionale olandese, Koninklijke Bibliotheek, gestisce un sistema per il deposito e la preservazione a lungo termine delle edizioni digitali, e-Depot,<sup>12</sup> che fa parte integrante della biblioteca stessa; la National Library of Australia sceglie e gestisce con il Progetto PANDORA,<sup>13</sup> l'archivio delle copie delle pubblicazioni australiane online significative, e dei siti web; la biblioteca nazionale tedesca insieme alla biblioteca dell'università di Goettingen, hanno realizzato l'archivio digitale a lungo termine KOPAL,<sup>14</sup> la Library of Congress ha lanciato il programma NDIIPP, National Digital Information Infrastructure and Preservation Programme<sup>15</sup>, volto a preservare i contenuti digitali dal rischio di perdita, in collaborazione con diverse biblioteche e organizzazioni. L'elemento che accomuna i diversi progetti elaborati al di fuori delle competenze delle biblioteche nazionali, invece, è la possibilità di consultare la letteratura prodotta in formato digitale nel caso di indisponibilità degli editori per svariati motivi, compresi gli eventi catastrofici e l'improvvisa chiusura di un'attività editoriale commerciale. A fianco di depositi digitali realizzati col sostegno degli editori, veri e propri *dark archives* cui le biblioteche potranno accedere solo nel caso di indisponibilità delle piattaforme editoriali, si sono sviluppati programmi cooperativi che mettono in grado le biblioteche stesse di preservare le copie digitali, ad esempio di e-journals, per le quali hanno sottoscritto con gli editori una licenza che consente l'auto-archiviazione e la consultazione perpetua, ma che offrono loro anche la possibilità di salvaguardare la produzione scientifica edita e diffusa nella modalità dell'accesso aperto.

Nell'aprile 2009 DOAJ Directory of open access journals<sup>16</sup> ha annunciato l'avvio della cooperazione con e-Depot della Koninklijke

---

<sup>12</sup><http://www.kb.nl/en/expertise/e-depot-and-digital-preservation>.

<sup>13</sup><http://pandora.nla.gov.au>

<sup>14</sup>[http://kopal.langzeitarchivierung.de/index\\_ziel.php.en](http://kopal.langzeitarchivierung.de/index_ziel.php.en).

<sup>15</sup><http://www.digitalpreservation.gov/about/background.html>.

<sup>16</sup>[www.doaj.org](http://www.doaj.org).

Bibliotheek-National Library of the Netherlands<sup>17</sup> per la preservazione a lungo termine delle riviste scientifiche open access, in modo da assicurarne la consultazione nel futuro. DOAJ è la banca dati internazionale, realizzata dalle biblioteche dell'Università di Lund in Svezia, che dal 2003 censisce le riviste in open access, che riconoscono quindi agli utenti il diritto di leggere, scaricare dalla rete, distribuire, stampare copie per il proprio uso di ricerca, creare link per scopi scientifici da questi articoli. Sono presenti più di 9.000 riviste di 120 nazioni, ricercabili anche secondo le aree disciplinari, con la possibilità di ottenere articoli in formato PDF.

La Koninklijke Bibliotheek, che coordina la European Alliance for Permanent Access, un'infrastruttura europea per l'accesso alle risorse informative, ha predisposto l'archiviazione delle pubblicazioni elettroniche in e-Depot, che assicura alla comunità dei ricercatori l'accesso perpetuo alle registrazioni pubblicate in tutti i campi scientifici, mentre i metadati sono registrati nel database bibliografico di KB. e-Depot, oltre a costituire l'archivio web olandese e delle immagini digitali, funziona come deposito nazionale delle collezioni elettroniche, ed ha poi esteso il servizio agli editori internazionali. Dal momento che e-Depot non è né un vero *dark archive* né un *light archive*, l'accesso viene stabilito da specifici accordi,<sup>18</sup> che tendono a bilanciare le esigenze del servizio agli utenti con gli interessi commerciali degli editori. L'accesso completo viene invece garantito solo per le pubblicazioni di editori che operano in open access, come Bio-Med Central<sup>19</sup>. Gli accordi realizzati dalla Koninklijke Bibliotheek prevedono infatti che sia permesso l'accesso a tutte le pubblicazioni archiviate dagli editori, tra i quali si annoverano Elsevier, Bio Med Central, Blackwell, Oxford University Press, Taylor & Francis,

<sup>17</sup>[www.kb.nl/e-depot-en](http://www.kb.nl/e-depot-en).

<sup>18</sup><http://www.doaj.org/doaj?func=news&nId=22&uiLanguage=en>.

<sup>19</sup><http://www.kb.nl/en/expertise/e-depot-and-digital-preservation/information-for-international-publishers>.

Springer, senza ostacolare i loro interessi commerciali, assicurando l'accesso permanente alle riviste a tutti gli utenti delle biblioteche autorizzate, inclusa la possibilità di document delivery interbibliotecario, limitato all'Olanda, e incluso l'accesso da remoto, se permesso. In caso di evento catastrofico, con indisponibilità dell'editore, l'archivio ufficiale entra a far parte del sistema che assicura il servizio, costituendo così una garanzia per tutte le biblioteche che hanno pagato una licenza (Oltmans e Van Wijngaarden). In specifiche condizioni, quando i titoli non sono più disponibili da parte dell'editore o di qualsiasi altra fonte, quando un editore cessa l'attività, oppure cessa la pubblicazione di una testata, o non offre più i *back issues*, ma anche in caso di catastrofi e danni alla piattaforma di distribuzione, l'accesso può essere consentito liberamente a chiunque tramite il sito web di e-Depot, senza considerare se le istituzioni avessero o no sottoscritto una licenza.<sup>20</sup>

Una recente indagine, tuttavia, ha messo in evidenza come il progetto di archiviazione in cooperazione tra DOAJ e e-Depot di KB – che, al momento, coinvolge le riviste di cui DOAJ gestisce la copia completa – comporti tempi di realizzazione piuttosto lunghi: circa 8 anni, tenendo conto che KB prevede l'archiviazione, orientativamente, di 1.560 riviste all'anno (Seadle). I tempi lunghi del progetto potrebbero comportare difficoltà soprattutto per le riviste scientifiche edite da piccoli istituti di ricerca, o da editori minori. Per questo motivo, quindi, sarebbe assolutamente necessario che le riviste censite da DOAJ fossero archiviate anche in altri depositi per la preservazione. La maggior parte delle riviste in open access presenti in DOAJ, in realtà, non sono invece presenti in depositi digitali per la preservazione e, come ha rilevato Michael Seadle, questo pone i documenti scientifici pubblicati in open access al rischio di

---

<sup>20</sup><http://www.kb.nl/en/expertise/e-depot-and-digital-preservation/more-about-the-e-depot>.

essere perduti. Alla fine del 2010 risultano poche, infatti, le riviste censite da DOAJ che siano trattate in archivi digitali come PORTICO oppure in LOCKSS/CLOCKSS: l'8% (445, di cui 356 in full text) in LOCKSS/CLOCKSS, e il 5% (288, di cui 270 in full text) in PORTICO (Seadle). Solo il 6% delle riviste full text presenti in DOAJ, risultano essere archiviate sia in PORTICO che in LOCKSS/CLOCKSS; la maggior parte di queste risultano sostanzialmente possedute da un esiguo numero di grandi editori, e questo significa che i piccoli editori sono più a rischio, dal momento che hanno maggiore difficoltà a sostenere i costi chiesti da PORTICO. La panoramica degli archivi digitali realizzati nel mondo per preservare la letteratura scientifica pubblicata su riviste elettroniche, è piuttosto nutrita, e contempla, oltre ai progetti posti in essere da biblioteche nazionali, realizzazioni in collaborazione tra diverse università, come MetaArchive, gestito in collaborazione tra le università della Florida, della Virginia, della Georgia, supportate dalla Library of Congress, per promuovere le pratiche di preservazione digitale distribuita, o come il progetto Chronopolis, realizzato dalla California Digital Library insieme al Inter-University Consortium for Political and Social Sciences, e infine, PubMedCentral, l'archivio della letteratura biomedica, gestito dalla National Library of Medicine.<sup>21</sup> Qui prendiamo ora in considerazione, entrando nei dettagli, le caratteristiche di due *dark archives*, uno dei quali realizzato con il contributo degli editori, e di un sistema in cooperazione posto in essere dalle biblioteche, in diversa misura rilevanti ai fini della preservazione della letteratura scientifica in open access.

---

<sup>21</sup>Una rassegna molto sintetica, e priva di dettagli, dei diversi progetti, è offerta da Gaur e Tripathi.

## 2.1 PORTICO

Proseguendo il progetto avviato sotto la direzione dell'Università del Michigan, nel 1995 la Andrew W. Mellon Foundation fonda JSTOR (Journal Storage), che dal 1997 mette a disposizione, dietro sottoscrizione, i fascicoli digitalizzati di annate pregresse di diverse riviste, allo scopo di agevolare la preservazione del materiale librario e migliorare le possibilità di ricerca degli studiosi, offrendo anche alle biblioteche e agli editori l'opportunità di conservare in formato digitale annate di riviste nate cartacee. Il modello non era adatto alla preservazione di riviste *digital born*, e infatti dal 2002 JSTOR ha attivato "The Electronic-Archiving Initiative", con l'obiettivo di realizzare un modello di archivio digitale per la letteratura scientifica edita in formato elettronico, e nel 2005 da avvio a PORTICO,<sup>22</sup> che si presenta come una soluzione per la preservazione dei documenti digitali di rilievo scientifico che viene incontro alle necessità degli editori e delle biblioteche, e successivamente sarà gestito dall'organizzazione senza scopo di lucro ITHAKA, e dal 2010 offre anche collezioni storiche digitalizzate ed e-books. PORTICO è uno dei più grandi archivi digitali al mondo, conforme al modello OAIS<sup>23</sup>, che preserva circa 15.000 e-journals, circa 177.000 e-books e documenti rilevanti per la comunità scientifica e accademica, affinché sia garantito agli studiosi l'accesso ad essi nel futuro. Il modello economico è sostenibile perché partners di PORTICO sono attualmente 217 editori (per l'Italia, Fabrizio Serra Editore e CLUEB) e 849 biblioteche affiliate. Il servizio di preservazione di e-journals si basa sulla collaborazione degli editori e delle biblioteche, per affrontare i costi della conser-

---

<sup>22</sup><http://www.portico.org>

<sup>23</sup>Reference model for an Open Archival Information System, realizzato nel 2002 da CCSDS-Consultative Committee for Space Data Systems, approvato nel 2003 come standard ISO 14721, e di cui nel giugno 2012 è stata rilasciata la *Recommendation for Space Data System Practices*: <http://public.ccsds.org/publications/archive/650x0m2.pdf>

vazione e della migrazione periodica dei dati. Nel caso di *trigger events*, e cioè nel caso in cui alcune testate di e-journals non siano più disponibili attraverso l'editore o altre fonti, ad esempio nel caso in cui un editore abbia cessato la pubblicazione, oppure non metta più a disposizione i fascicoli precedentemente pubblicati, o nel caso di eventi catastrofici che coinvolgono la piattaforma distributiva dell'editore, l'accesso è riservato alle biblioteche che partecipano all'iniziativa. In questi casi le riviste sono disponibili attraverso il sito web dell'archivio per tutte le biblioteche che partecipano, senza considerare se le istituzioni partecipanti avevano o meno attivato una licenza. Dal 2008 è stato avviato il servizio di preservazione di e-books per gli editori, allargato dal 2011 anche alle biblioteche, con le medesime caratteristiche di quello per gli e-journals.

Le condizioni per l'uso del *dark archive* prevedono che solo nei casi di *trigger events* o di catastrofe, le istituzioni e le biblioteche che partecipano, e i loro utenti autorizzati, possano accedere ai documenti, dopo un periodo di più di 90 giorni di mancanza di possibilità di accesso; le biblioteche in ogni caso non potranno usare l'accesso per distribuire i files ad altri che non siano autorizzati, non potranno attuare il prestito interbibliotecario, o il downloading, o utilizzare crawler e spider per esportare i contenuti<sup>24</sup>. Gli editori, che contribuiscono finanziariamente all'impresa, sottoscrivono una licenza non esclusiva, che consente a PORTICO di gestire e archiviare i files, normalizzandone i formati – attività non esente dal rischio di danni del contenuto originale – e permette di adottare le procedure di migrazione dei dati; le biblioteche contribuiscono col pagamento di una iscrizione annuale e forniscono gli identificativi per l'autenticazione degli utenti (Galyani Moghaddam, p.87).

È importante sottolineare il fatto che questo archivio dal 2008 ha incrementato la preservazione delle riviste edite in open access, e

---

<sup>24</sup><http://www.portico.org/digital-preservation/terms-of-use>.

attualmente archivia circa 400 riviste, tra cui l'intera collezione di 60.000 articoli *peer-reviewed* pubblicati in open access da BioMed Central, le riviste editate da Bentham Open e, recentemente, dall'italiana PAGEPress.<sup>25</sup> In base alla politica per l'utilizzazione dell'archivio, a seguito di un *trigger event*, anche queste pubblicazioni saranno disponibili soltanto alle biblioteche che partecipano all'iniziativa, e ai loro utenti identificati, vanificando in questo modo, in realtà, la filosofia dell'accesso aperto e contravvenendo alle decisioni degli autori, che in origine avevano scelto questo canale di diffusione dei risultati delle ricerche.

## 2.2 LOCKSS

Progettato a partire dal 1999 presso la Stanford University, e realizzato nella versione definitiva nel 2001, Lots of Copies Keep Stuff Safe (LOCKSS),<sup>26</sup> è un software che permette alle istituzioni bibliotecarie di archiviare e preservare a basso costo le collezioni digitali per le quali avevano pagato la sottoscrizione, senza dover fare ricorso ad agenzie o strutture esterne, ma attuando invece una prassi cooperativa. L'iniziativa è stata sostenuta finanziariamente da Andrew W. Mellon Foundation, National Science Foundation e Library of Congress, e il sistema viene attualmente finanziato dalle biblioteche che partecipano all'iniziativa. L'uso del software open source LOCKSS permette alle biblioteche di mantenere in repository controllati le collezioni digitali *e-only*, per le quali esse hanno sottoscritto una licenza con gli editori per l'accesso "in perpetuo", ma anche documenti editi in open access. I contenuti preservati diventano parte della collezione della biblioteca, che potrà controllarne continuamente

---

<sup>25</sup>Come esempio, si può consultare: <http://www.portico.org/digital-preservation/news-events/news/newly-signed-publishers/portico-to-preserve-open-access-content-from-biomed-central>

<sup>26</sup><http://lockss.stanford.edu/lockss/Home>

l'integrità. Le biblioteche depositano infatti ciascuna nella propria LOCKSS box, un computer dedicato allo scopo, la copia digitale acquisita attraverso crawler direttamente dai siti web degli editori, che devono avere accordato il permesso, segnalandolo con una dichiarazione sul proprio sito web. Il sistema permette alle biblioteche stesse di preservare a lungo termine e consentire l'accesso per la fruizione ai propri utenti nei casi in cui il server dell'editore, che avrebbe dovuto garantire l'accesso perpetuo, sia momentaneamente indisponibile per qualsiasi motivo, sia danneggiato, sotto attacco di *hacker*, o nel caso in cui una pubblicazione non sia più online, o l'editore abbia addirittura cessato l'attività. Con questo sistema, le biblioteche possono utilizzare in sostanza con i documenti *e-only* le modalità che usavano con i libri cartacei, i quali, una volta acquistati, entravano a far parte del patrimonio della biblioteca, che ne poteva disporre in ogni caso. In tal modo, vengono rispettati anche le esigenze e i diritti degli editori commerciali, che desiderano avere la certezza che i loro prodotti editoriali siano accessibili soltanto a chi è autorizzato. Il timore degli editori, che i contenuti digitali possano essere replicati illegalmente se custoditi da altri, con questo sistema viene facilmente superato, dal momento che i contenuti possono essere forniti ad altre LOCKSS boxes solo nel caso in cui si debbano sostituire contenuti già posseduti e risultati danneggiati. I 520 editori che partecipano all'iniziativa, tra i quali Springer, Taylor & Francis, e molte University Press, godono poi del vantaggio di poter delegare alle biblioteche la responsabilità e i costi della preservazione a lungo termine.

I documenti digitali conservati in una LOCKSS box sono continuamente monitorati e viene verificata la loro integrità, attraverso una comparazione con i medesimi documenti conservati in altre LOCKSS boxes, selezionate a caso. Nel caso in cui un contenuto risulti danneggiato, o non sia completo, sarà ripristinato grazie alla collaborazione

delle altre biblioteche, che costituiscono una rete di preservazione digitale decentralizzata, peer-to-peer. Lo staff di ciascuna biblioteca controlla l'accesso alle riviste preservate (Dobson; Reich e Rosenthal, «LOCKSS: A Permanent Web Publishing and Access System»; Reich e Rosenthal, «Distributed Digital Preservation: Private LOCKSS Networks and Business, Social, and Technical Frameworks»). Il processo di controllo del contenuto memorizzato è completamente automatico: i contenuti di una LOCKSS box vengono continuamente comparati con quelli conservati nelle memorie delle altre boxes.

Il processo viene realizzato tramite il protocollo peer-to-peer LCAP (Library Cache Auditing Protocol), e nel caso in cui siano riscontrate differenze tra due copie del medesimo contenuto, verrà acquisita automaticamente, senza l'intervento umano, una nuova copia o dal sito dell'editore, o da altra box. LOCKSS affronta anche il problema dell'obsolescenza dei formati predisponendo la "migrazione su richiesta". I documenti sono convertiti in un nuovo formato leggibile solo quando un utente lo richiede, e comunque viene preservato il più possibile il formato originale. Il sistema è adatto ad archiviare contenuti stabili, che non vengono modificati con frequenza, ad esempio una copia integrale di una rivista elettronica, ma non un sito web (Prudlo; Schneider).

Global LOCKSS Network è la rete istituita dalle LOCKSS boxes create dalle biblioteche per preservare le collezioni di e-journals e di e-books per i quali esse avevano sottoscritto una licenza d'uso, e le collezioni di documenti che sono stati editi in open access. Dal 2004 è richiesto alle biblioteche che partecipano un modesto contributo per l'iscrizione, in cambio del quale esse entrano nella LOCKSS Alliance, ricevono il supporto tecnico da parte dello staff della Stanford University per creare e gestire il loro deposito, e l'accesso perpetuo alle collezioni. Le LOCKSS boxes si configurano come veri e propri scaffali digitali per preservare le collezioni utilizzando la tecnologia

LOCKSS per acquisire la proprietà dei contenuti, costruire collezioni permanenti, di cui le biblioteche avranno l'accesso perpetuo, invece di dover contare solo su di un accesso temporaneo attraverso una licenza. Sono le biblioteche a scegliere quali contenuti preservare, e l'archivio resta sotto il loro controllo. Bisogna sottolineare che in questo caso si permette alle biblioteche, ad esempio a quelle accademiche, che abbiano selezionato interi numeri di riviste scientifiche in open access, o singoli articoli, di preservarli contro qualsiasi evento possa in futuro impedirne la fruizione: ad esempio, nel caso in cui la consultazione di qualche annata di una rivista scientifica, edita da una Università o da una associazione scientifica o professionale, non venga più garantita in accesso aperto per banali motivi di gestione. Private LOCKSS Networks costituisce una rete cui partecipano biblioteche che possiedono collezioni altamente specializzate o particolarmente significative: immagini, manoscritti, documenti governativi, documenti audiovisivi, che saranno resi accessibili solo all'istituzione cui appartengono, e solo in caso di necessità di recuperare contenuti perduti a causa di disastri naturali o incidenti. Ciascun PLN generalmente raggruppa da sette a dodici istituzioni omogenee per finalità e scopi, e decide la propria *policy* per l'accesso, dipendente dalle tipologie di contenuti e dalle specifiche esigenze. Un ulteriore progetto, Open LOCKSS, esclusivamente dedicato alla preservazione delle riviste scientifiche ad accesso aperto, lanciato nel 2007 nel Regno Unito dal Joint Information Systems Committee (JISC), è gestito dall'Università di Glasgow. Prevede l'inclusione in LOCKSS delle riviste editate da piccoli gruppi editoriali, o da società scientifiche e dipartimenti universitari, che hanno adottato la pubblicazione in accesso aperto, e che non possiedono risorse sufficienti per partecipare ad archivi per la preservazione a lungo termine.<sup>27</sup> Nella prima fase, furono inseriti nell'archivio LOCKSS diciannove

---

<sup>27</sup><http://www.lib.gla.ac.uk/Research/openlockss/forpublishers.shtml>.

periodici.<sup>28</sup>

## 2.3 CLOCKSS

CLOCKSS (Controlled LOCKSS),<sup>29</sup> attivo dal 2006, è una organizzazione no-profit di editori che operano nel settore accademico e di biblioteche di ricerca – tra le quali, Australian National University, Stanford University, University of Virginia, ma anche OCLC e New York Public Library – volta a realizzare una rete distribuita, a livello mondiale, di *dark archives* che utilizzano la tecnologia LOCKSS, allo scopo di preservare a lungo termine le pubblicazioni scientifiche in formato digitale, in particolare gli e-journals. Gli editori che partecipano – tra i quali, Elsevier, Springer, Taylor & Francis, Wiley-Blackwell e diverse University Press, e per l'Italia Casalini Libri, Fabrizio Serra, Cadmo e CLUEB – depositano i loro *front* e *back files*, per la conservazione permanente. I documenti digitali, dei quali gli editori eventualmente avessero deciso di sospendere la pubblicazione, così come i documenti pubblicati in open access ma non più disponibili su alcuna piattaforma, possono essere messi a libera disposizione, a seguito della decisione del Board of Directors and Advisory Council.

CLOCKSS è un sistema chiuso, utilizzato dalle biblioteche e dagli editori come *dark archive* sicuro e dai costi di gestione limitati, e creato allo scopo di preservare i contenuti digitali per tutte le biblioteche del mondo. Nel caso di eventi eccezionali che causano l'indisponibilità dei documenti *e-only* pubblicati da editori commerciali, i *trigger event*, ad esempio nel caso in cui l'editore non sia più in grado di dare accesso a documenti pubblicati in precedenza, per motivi tecnici, o per cause dovute a catastrofi naturali, oppure, per

---

<sup>28</sup>Cfr. il Final Report dell'agosto 2007, [http://www.lib.gla.ac.uk/Research/openlockss/Openlockss\\_%20Final\\_Report\\_Public\\_%20Aug08.pdf](http://www.lib.gla.ac.uk/Research/openlockss/Openlockss_%20Final_Report_Public_%20Aug08.pdf)

<sup>29</sup><http://www.clockss.org/clockss/Home>.

una sua decisione di politica editoriale cessi la pubblicazione di un periodico, ma anche nel caso in cui un editore cessi l'attività senza che vi siano eredi diretti in grado di continuare senza interruzione l'erogazione dei servizi, il sistema interviene ripristinando la consultazione dei documenti senza alcun costo né per le biblioteche che lo finanziano, né per tutti coloro che avevano sottoscritto una risorsa in abbonamento in precedenza. La soluzione di *dark archive* posta in essere da CLOCKSS prevede che quindici biblioteche dislocate in aree geografiche diverse, dall'America del nord all'Europa, all'Asia, dalle situazioni geologiche e politiche più disparate, conservino una copia completa dell'archivio.

Le biblioteche contribuiscono finanziariamente a sostenere CLOCKSS per un periodo di cinque anni, dopo il quale l'impegno economico diminuisce, e a fronte di ciò cooperano insieme agli editori nello stabilire le politiche gestionali del *dark archive*. Il sostegno finanziario è necessario per poter erogare tutti i servizi, rinforzare il circuito, attraendo editori minori, le cui pubblicazioni potrebbero correre maggiori rischi. Le politiche adottate per la consultazione dei contenuti offerti ad accesso aperto, costituiscono un elemento critico, di fondamentale importanza per la valutazione di questi archivi. Se e-Depot, gestito dalla Koninklijke Bibliotheek, garantisce a chiunque la consultazione completa di tutta la produzione editoriale realizzata nell'ambito della filosofia dell'open access, e CLOCKSS rende disponibili i contenuti open access dopo il verificarsi di un *trigger event*, quando ad esempio non siano più curati dall'ente editore,<sup>30</sup> oltre a disporre in libera consultazione, nel caso in cui si verifichi il medesimo evento scatenante, i contenuti in precedenza fruibili a pagamento, PORTICO adotta invece la medesima politica "after trigger" sia per le riviste che erano consultabili a seguito del pa-

---

<sup>30</sup>"Open access content no longer available from the publisher, is available for free from CLOCKSS" <http://www.clockss.org/clockss/Home>

gamento di una licenza d'uso, sia per le riviste ad accesso aperto, restringendo la consultazione anche dei contenuti in open access alle biblioteche che partecipano all'iniziativa.

### 3 Conclusioni

La salvaguardia delle pubblicazioni scientificamente rilevanti, edite in open access, e non più disponibili liberamente, in quanto abbandonate dagli enti editori, sta diventando una necessità inderogabile per preservare la possibilità per i ricercatori di accedere ai contenuti e ai risultati della ricerca scientifica e utilizzarli per ulteriori ricerche. I sistemi LOCKSS e CLOCKSS possono servire egregiamente a questo scopo. Per quanto riguarda l'Italia, oltre agli editori prima citati, partecipano a CLOCKSS l'Istituto Nazionale di Geofisica e Vulcanologia (INGV), la Libera Università Internazionale degli Studi Sociali (LUISS), il Politecnico di Milano, le Università degli studi di Bergamo, Brescia, Ferrara, Milano, Verona, Napoli, e l'Istituto Universitario Europeo, con le loro biblioteche. L'Università Cattolica del Sacro Cuore di Milano, oltre a partecipare a LOCKSS (D'Amato e Mazzocchi) dal 2010 costituisce uno dei nodi di CLOCKSS. Sarebbe opportuno che tutte le biblioteche degli Atenei italiani, e degli enti di ricerca, si attivassero per entrare a far parte di questi circuiti, almeno della rete cooperativa LOCKSS, per preservare le pubblicazioni elettroniche acquisite dagli editori commerciali, e quelle disponibili in modalità open access.

### Riferimenti bibliografici

Antelman, Kristin. «Do Open-Access Articles Have a Greater Research Impact?» *College & Research Libraries* 65.5. (2004): 372–382. <<http://crl.acrl.org/content/65/5/372.full.pdf+html>>. (Cit. a p. 129).

- Bailey, Charles W., Jr. *Open Access Bibliography: Liberating Scholarly Literature with E-Prints and Open Access Journals*. Houston, Tx: Digital Scholarship. <http://digital-scholarship.org/oab/oab.htm>. (Cit. a p. 122).
- Bergamin, Giovanni e Maurizio Messina. «Magazzini digitali: dal prototipo al servizio». *Digitalia* 5.2. (2010): 144–153. (Cit. a p. 131).
- Björk, Bo-Christer. «Open Access to the Scientific Journal Literature: Situation 2009». *PLoS ONE* 5.6. (2010). DOI: [10.1371/journal.pone.0011273](https://doi.org/10.1371/journal.pone.0011273). (Cit. a p. 125).
- Cassella, Maria. «Le biblioteche nella transizione al "solo elettronico"». *Biblioteche oggi* 26.5. (2008): 3–13. <<http://www.bibliotecheoggi.it/2008/20080500301.pdf>>. (Cit. a p. 131).
- . *Open access e comunicazione scientifica. Verso un nuovo modello di disseminazione della conoscenza*. Milano: Editrice Bibliografica, 2012. (Cit. a p. 123).
- Cullen, Rowena e Brenda Chawner. «Institutional Repositories, Open Access, and Scholarly Communication: A Study of Conflicting Paradigms». *The Journal of Academic Librarianship* 37.6. (2011): 460–470. <<http://www.sciencedirect.com/science/article/pii/S009913331100156X>>. (Cit. alle pp. 126, 127).
- D'Amato, Gianluca e Renato Mazzocchi. «Conservando più copie rendiamo sicure le nostre collezioni digitali. La sperimentazione della tecnologia LOCKSS presso il Sistema bibliotecario dell'Università Cattolica del Sacro Cuore». *Biblioteche oggi* 7. (2008): 9–13. <<http://www.bibliotecheoggi.it/2008/20080700901.pdf>>. (Cit. a p. 144).
- De Robbio, Antonella. «Accesso aperto 2012: la vie en rose». *AIB Studi* 52.1. DOI: [10.2426/aibstudi-6293](https://doi.org/10.2426/aibstudi-6293). (2012): 17–29.
- . «Analisi citazionale e indicatori bibliometrici nel modello Open Access». *Bollettino AIB* 47.3. (2007): 257–289. <<http://bollettino.aib.it/article/view/5246/5013>>. (Cit. a p. 128).
- Dobson, Chris. «From bright idea to beta test. The story of LOCKSS». *Searcher. The Magazine for Database Professionals* 11.2. (2003): 50–53. <<http://www.infoday.com/searcher/feb03/index.shtml>>. (Cit. a p. 140).
- Galyani Moghaddam, Golnessa. «Preserving scientific electronic journals: a study of archiving initiatives». *The electronic library* 26.1. (2008): 83–96. DOI: [10.1108/02640470810851761](https://doi.org/10.1108/02640470810851761). (Cit. a p. 137).
- Gargouri, Yassine, et al. «Self-Selected or Mandates, One Access Increases Citation Impact for Higher Quality Research». *PLoS ONE* 5.10. (2010). DOI: [10.1371/journal.pone.0013636](https://doi.org/10.1371/journal.pone.0013636). (Cit. a p. 130).
- Gaur, Ramesh C. e Manorama Tripathi. «Digital Preservation of Electronic Resources». *DESIDOC Journal of Library & Information Technology* 32.4. (2012): 293–301. <<http://www.digitalmeetsculture.net/wp-content/uploads/2012/11/Gaur-article.pdf>>. (Cit. a p. 135).

- Giordano, Tommaso. «Le collezioni non abitano più qui? Conservazione e strategie di cooperazione in transizione». *Biblioteche oggi* 2. (2006): 90–102. <<http://www.bibliotecheoggi.it/2006/20060209001.pdf>>. (Cit. a p. 131).
- Guerrini, Mauro. *Gli archivi istituzionali. Open access, valutazione della ricerca e diritto d'autore*. A cura di Andrea Capaccioni. Milano: Editrice Bibliografica, 2010. (Cit. a p. 123).
- Hajjem, Chawki, Stevan Harnad e Yves Gringras. «Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact». *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*. (2005). <<http://arxiv.org/ftp/cs/papers/0606/0606079.pdf>>. (Cit. a p. 129).
- Kirchoff, Amy e Eileen Gifford Fenton. «Archiviazione di periodici elettronici. L'approccio di Portico, un progetto nato negli Stati Uniti». *Biblioteche oggi* 9. (2007): 7–15. <<http://www.bibliotecheoggi.it/2007/20070900701.pdf>>. (Cit. a p. 131).
- McVeigh, Marie E. «Open Access Journals in the ISI Citation Databases: Analysis of Impact Factors and Citation Patterns. A citation study from Thomson Scientific». (2004). <<http://ip-science.thomsonreuters.com/m/pdfs/openaccessitions2.pdf>>. (Cit. a p. 125).
- Norris, Micheal, Charles Oppenheim e Fytton Rowland. «The Citation Advantage of Open-Access Articles». *Journal of the American Society for Information Science and Technology* 59.12. (2008): 1963–1972. <<http://onlinelibrary.wiley.com/doi/10.1002/asi.20898/full>>. (Cit. a p. 130).
- Oltmans, Erik e Hilde Van Wijngaarden. «The KB e-Depot digital archiving policy». *Library Hi Tech* 24.4. DOI: [10.1108/07378830610715464](https://doi.org/10.1108/07378830610715464). (2006). (Cit. a p. 134).
- Prudlo, Marion. «E-Archiving: An Overview of Some Repository Management Software Tools». *Ariadne* 43. (2005). <<http://www.ariadne.ac.uk/print/issue43/prudlo>>. (Cit. a p. 140).
- Reich, Vicky e David S. H. Rosenthal. «LOCKSS: A Permanent Web Publishing and Access System». *D-Lib Magazine* 7.6. (2001). <<http://www.dlib.org/dlib/june01/reich/06reich.html>>. (Cit. a p. 140).
- Reich, Victoria e David Rosenthal. «Distributed Digital Preservation: Private LOCKSS Networks and Business, Social, and Technical Frameworks». *Library Trends* 57.3. DOI: [10.1353/lib.0.0047](https://doi.org/10.1353/lib.0.0047). (2009): 461–475. (Cit. a p. 140).
- Schneider, Karen G. «Lots of Librarians Can Keep Stuff Safe». *Library Journal* August. (2007): 30–31. <<http://lj.libraryjournal.com/2007/08/ljarchives/lots-of-librarians-can-keep-stuff-safe/>>. (Cit. a p. 140).
- Seadle, Micheal. «Archiving in the networked world: open access journals». *Library Hi Tech* 29.2. DOI: [10.1108/0737883111138251](https://doi.org/10.1108/0737883111138251). (2011): 394–404. (Cit. alle pp. 134, 135).

- Waters, Donald J., cur. *Urgent Action Needed to Preserve Scholarly Electronic Journals*. New York: The Andrew W. Mellon Foundation. <http://www.arl.org/publications-resources/1150>. (Cit. alle pp. 121, 122).
- Way, Doug. «The Open Access Availability of Library and Information Science Literature». *College & Research Libraries* 71.4. (2010): 302–309. <<http://crl.acrl.org/content/71/4/302.full.pdf+html>>. (Cit. a p. 126).
- Xia, Jingfeng. «Diffusionism and open access». *Journal of Documentation* 68.1. DOI: [10.1108/07378831111138251](https://doi.org/10.1108/07378831111138251). (2012): 72–98. (Cit. alle pp. 124, 125).

MARIA TERESA BIAGETTI, Università La Sapienza.

[mariateresa.biagetti@uniroma1.it](mailto:mariateresa.biagetti@uniroma1.it)

---

Biagetti, M. T. "Preservare la letteratura scientifica elettronica in open access". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #9025. DOI: [10.4403/jlis.it-9025](https://doi.org/10.4403/jlis.it-9025). Web.

**ABSTRACT:** There has been a remarkable growth of the use of e-resources, and especially of e-journals by researchers, in every field of studies. Ensuring the accessibility and use by future generations is an important challenge for university and research libraries. The paper is particularly dealing with the problems of digital preservation of relevant scientific documents issued according to the Open access movement. After examining the characteristics of the diffusion and citation impact of Open access papers in scientific literature, the paper highlights the need for all the research libraries to preserve digital contents published in OA, to ensure the access and use for future researchers. The KB e-Depot and DOAJ project for OA digital preservation is taken into consideration, and systems for digital archiving and for cooperative preservation of scholarly journals' contents: PORTICO, the LOCKSS Alliance and CLOCKSS, are presented.

**KEYWORDS:** Digital preservation; Open Access; Scholarly Journals; PORTICO; LOCKSS; CLOCKSS.

---

Submission: 2013-08-14

Accettazione: 2013-09-28

Pubblicazione: 2014-01-01





# FLORe – Florence Repository. L'Archivio ad accesso aperto dell'Università di Firenze

Gaia Innocenti, Cristina Mugnai, Valdo Pasqui

## 1 Introduzione

Anche in Italia l'accesso aperto alle pubblicazioni scientifiche è diventato ormai un tema di pressante attualità sia dal punto di vista normativo che per quanto riguarda gli aspetti organizzativi e le ricadute operative nel mondo della ricerca. L'open access non è più, come in anni ancora recenti, una chimera, la visione dei bibliotecari e dei ricercatori di alcuni ambiti disciplinari, ma si tratta di una realtà in crescita sorretta dalla maggiore consapevolezza dei vantaggi che ne derivano da parte degli attori coinvolti e supportata da linee d'indirizzo che insieme concorrono a realizzare uno scenario impensabile e insperato fino a pochi anni fa. La Commissione Europea ha sancito che l'accesso aperto alle pubblicazioni scientifiche è uno dei principi del nuovo Framework Programme for Research and Innovation, Horizon 2020, la cui portata si estende dal 2014 al 2020 e nel quale è stato stabilito che a partire dal 2014 tutti gli articoli prodotti grazie al finanziamento di fondi del programma dovranno essere accessibili secondo una delle due modalità consolidate:



- a. "gold" open access—articoli immediatamente resi accessibili online dall'editore;
- b. "green" open access—i ricercatori rendono i loro articoli disponibili attraverso repository ad accesso aperto entro 6 mesi (12 per gli articoli negli ambiti delle scienze sociali e umanistiche).

In un comunicato stampa del 17 luglio 2013 la Commissione Europea ha ribadito questa impostazione raccomandando agli «Stati membri di adottare un approccio simile anche per i risultati della ricerca finanziata con i programmi nazionali. L'obiettivo è rendere accessibile entro il 2016 il 60% degli articoli scientifici su lavori finanziati con fondi pubblici europei secondo la modalità dell'accesso aperto» e inoltre la Commissione ha espresso la volontà di iniziare anche a «sperimentare l'accesso aperto con i dati raccolti nel corso delle ricerche finanziate con fondi pubblici».<sup>1</sup> Nel mese di agosto 2013 è stato pubblicato uno studio (Archambault et al.) condotto da Science-Metrix per conto della Direzione Generale Ricerca & Innovazione della Commissione Europea che ha esaminato 22 ambiti di conoscenza su un campione di paesi comprendente, oltre ai 28 dell'Area Europea, anche Svizzera, Norvegia, Islanda, Brasile, Canada, Giappone, Stati Uniti ed altri. Lo studio afferma che in molti di questi paesi e per parecchie aree disciplinari ormai più del 50% degli articoli sono disponibili gratuitamente e che addirittura in discipline come la tecnologia, la ricerca biomedica, la biologia, la matematica e la statistica la maggioranza dei lavori sono disponibili ad accesso libero. In Italia, dopo le azioni di sensibilizzazione condotte dalla Commissione Open Access della CRUI<sup>2</sup> diversi Atenei hanno inserito nei propri statuti impegni precisi rispetto alla promozione dell'accesso aperto. Fatto del tutto nuovo ed in parte imprevisto, il 10 agosto è entrato in vigore il Decreto Legge 8

---

<sup>1</sup>[http://europa.eu/rapid/press-release\\_IP-12-790\\_it.htm](http://europa.eu/rapid/press-release_IP-12-790_it.htm).

<sup>2</sup><http://www.cru.it/HomePage.aspx?ref=894>.

agosto 2013, n. 91, recante "Disposizioni urgenti per la tutela, la valorizzazione e il rilancio dei beni e delle attività culturali e del turismo" il cui Art.4 comma 2 prevedeva che «Le pubblicazioni che documentano i risultati di ricerche finanziate per una quota pari o superiore al cinquanta per cento con fondi pubblici, indipendentemente dal formato della prima pubblicazione e dalle modalità della sua distribuzione o messa a disposizione del pubblico, devono essere depositate, non oltre sei mesi dalla pubblicazione, in archivi elettronici istituzionali o di settore, predisposti in modo tale da garantire l'accesso aperto, libero e gratuito, dal luogo e nel momento scelti individualmente, l'interoperabilità all'interno e all'esterno dell'Unione Europea e la conservazione a lungo termine in formato elettronico». Successivamente, in sede di conversione in legge, la Commissione cultura del Senato ha ampiamente emendato il testo attenuandone molto il dispositivo iniziale. Infatti la formulazione finale, approvata dai due rami del Parlamento, recita: «I soggetti pubblici preposti all'erogazione o alla gestione dei finanziamenti della ricerca scientifica adottano, nella loro autonomia, le misure necessarie per la promozione dell'accesso aperto ai risultati della ricerca finanziata per una quota pari o superiore al cinquanta per cento con fondi pubblici, quando documentate in articoli pubblicati su periodici a carattere scientifico che abbiano almeno due uscite annue. L'accesso aperto si realizza: a) tramite la pubblicazione da parte dell'editore, al momento della prima pubblicazione, in modo tale che l'articolo sia accessibile a titolo gratuito dal luogo e nel momento scelti individualmente; b) tramite la ripubblicazione senza fini di lucro in archivi elettronici istituzionali o disciplinari, secondo le stesse modalità, entro 18 mesi dalla prima pubblicazione per le pubblicazioni delle aree disciplinari scientifico-tecnico-mediche e 24 mesi per le aree disciplinari umanistiche e delle scienze sociali». Pur nella profonda revisione del contenuto la disposizione convertita in

legge deve essere accolta come il segnale che l'Italia ha comunque iniziato a recepire dal punto di vista normativo gli orientamenti europei e che è pronta a percorrere in modo più formalizzato e concreto la strada dell'accesso aperto. In questo contesto l'Università di Firenze è stata un'antesignana poiché già nell'anno 2000, tra i primi atenei in Italia, ha avviato la sperimentazione di un repository open access promuovendo il "self archiving" ("green road") con un progetto all'epoca pionieristico perché nato senza un quadro di riferimento tecnologico, bibliografico (i metadati), organizzativo e normativo consolidato e pertanto volto anche a valutare e sperimentare sia le possibili soluzioni tecniche che i modelli organizzativi e di comunicazione più adatti per promuovere la "cultura" dell'accesso aperto.

## 2 L'open access all'Università di Firenze

### 2.1 L'esperienza pregressa

Il primo archivio open access dell'ateneo fiorentino, denominato Archivio E-Prints, è stato realizzato oltre 10 anni fa nell'ambito del progetto per la costituzione della biblioteca digitale promosso dalla Firenze University Press (FUP). In via preliminare uno studio svolto nella primavera del 2000 (Pasqui) permise di delineare il framework di riferimento articolato, sulla base delle visioni e delle prospettive di quegli anni, secondo tre principi:

- a. preservare l'autonomia, sia dal punto di vista gestionale che da quello dei contenuti, delle iniziative esistenti e future di facoltà, dipartimenti e centri;
- b. garantire l'interoperabilità tra gli archivi in corso di costituzione;

- c. consentire agli utenti finali l'accesso integrato alle risorse, in particolare al catalogo in linea delle biblioteche (OPAC), alle collezioni digitali primarie e ai *subject information gateways*.

Tenendo conto di questi principi e del contesto funzionale (tipi utente, tipologie dei materiali, modalità di auto-archiviazione) lo studio condusse all'individuazione di un modello basato sulla Open Archives Initiative (OAI) assumendo in particolare l'articolazione in Data Provider e Service Provider e il protocollo Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Per la predisposizione dell'archivio, tra i sw open source all'epoca disponibili, fu scelto il software eprints.org sviluppato dall'Università di Southampton. L'archivio fu avviato nel gennaio 2001. Dopo una fase di sperimentazione, durata circa un biennio, il repository raggiunse uno stadio di consolidamento attraverso alcuni interventi di personalizzazione e configurazione del sw eprints.org quali lo sviluppo di una versione bilingue (italiano/inglese) e l'implementazione dello schema di soggettazione conforme ai settori scientifico disciplinari (14 Aree) definiti dal Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR)<sup>3</sup>. La registrazione presso l'harvester mondiale OAIster garantiva la possibilità di promuovere la disseminazione della produzione scientifica Open Access (OA) dell'ateneo. Dal punto di vista organizzativo e operativo le politiche di governo dei contenuti e dei rapporti con gli autori erano gestite dalla FUP, mentre il Centro Servizi Informatici dell'Ateneo Fiorentino (ora Sistema Informatico dell'Ateneo Fiorentino (SIAF)) garantiva lo sviluppo e la gestione tecnologica. Nel febbraio del 2004, in occasione del convegno organizzato per la presentazione ufficiale dell'archivio, alcuni interventi misero in luce aspetti che a distanza di anni si sarebbero rivelati premonitori delle criticità manifestatesi per questa iniziativa. Giancarlo Pepeu, allora Presidente della FUP, nella introduzione

---

<sup>3</sup><http://www.miur.it/UserFiles/115.htm>.

afferitava che «qualsiasi archivio e-prints deve dare garanzie sulla correttezza e l'attendibilità di quanto è in esso contenuto» e «se non vi è una garanzia della validità di quanto si trova negli archivi e-prints, essi diventeranno depositi di informazioni dubbie, di scarso valore, ai quali nessuno attingerà» (Pepeu). Patrizia Cotoneschi, nel porre in risalto le sinergie dell'archivio con l'attività editoriale, sottolineava che l'archiviazione non è «un'alternativa alla pubblicazione tradizionale, ma un mezzo complementare per diffondere i risultati della ricerca» e annunciava la predisposizione di un nuovo tipo di contratto FUP che avrebbe contemplato per gli autori «la possibilità di diffondere liberamente il formato elettronico delle loro pubblicazioni» (Cotoneschi). Infine, Giulia Maraviglia, dopo aver ricordato la consistenza dell'archivio (116 contributi per un totale di 56 autori), rilevava che la

oggettiva lentezza con cui sta crescendo l'archivio e-print dell'ateneo fiorentino può essere quasi certamente attribuita alla scarsa fiducia che la comunità accademica dimostra ancora di avere verso questa iniziativa. Ai bibliotecari tuttavia spetta il compito di promuovere e diffondere l'utilizzo di questo strumento, spesso per molti ancora sconosciuto. L'auspicio è che il deposito e la messa a disposizione dei materiali sopra elencati diventi prassi di lavoro in tutti gli atenei (Maraviglia).

Nonostante l'Università di Firenze il 6 giugno 2005 avesse aderito alla Dichiarazione di Berlino per l'accesso aperto alla letteratura scientifica<sup>4</sup>, negli anni successivi è venuta meno la capacità di intraprendere azioni in grado di promuovere efficacemente il deposito della produzione scientifica nell'archivio e di assicurarne uno sviluppo sistematico e coerente. Per queste criticità è stata determinante

---

<sup>4</sup><http://openaccess.mpg.de/286432/Berlin-Declaration>.

soprattutto l'assenza di una policy che definisse gli attori, i principi e gli obblighi, fatto che ha fortemente limitato lo sviluppo del self-archiving. Di conseguenza l'archivio è rimasto uno strumento di nicchia per pochi autori, con un contenuto molto eterogeneo concernente non solo la produzione scientifica, incluse alcune tesi di dottorato, ma anche supporti didattici di vario tipo (dispense e slide) ed altro materiale per una consistenza che alla fine del 2011 superava di poco i 1200 documenti, con un numero di nuovi contributi per anno inferiore a 100 eccetto pochi anni (2004, 2005 e 2007). Ovviamente il limitato utilizzo dell'archivio e-prints ne ha determinato anche l'obsolescenza tecnologica, pur non inficiandone il funzionamento. Infatti è stato inevitabile dedicare le risorse umane disponibili in possesso di competenze informatiche ai progetti dell'ateneo con più elevata priorità, spesso finanziati con contributi esterni (fondazioni bancarie, MIUR, Presidenza del Consiglio dei Ministri), relativi ai servizi online per studenti (es. la verbalizzazione degli esami con la firma digitale, la valutazione della didattica, il fascicolo elettronico dello studente etc). Nell'ultimo biennio, come vedremo tra breve, il riesame delle prospettive e la rivalutazione delle strategie relative all'accesso aperto all'interno dell'Università di Firenze hanno portato a sviluppare e realizzare una nuova soluzione per il repository istituzionale dell'ateneo.

## 2.2 La genesi di Florence Repository (FLORe)

Sulla base dall'esperienza pionieristica, ma dagli esiti non soddisfacenti, dell'archivio e-prints descritta nel paragrafo precedente e grazie all'opera di advocacy a favore dell'accesso aperto condotta da Mauro Guerrini nel maggio del 2011 il rettore dell'Università Firenze ha istituito la commissione di ateneo "sul tema accesso aperto per i prodotti della ricerca e per le tesi di dottorato", presieduta dal prorettore alla ricerca Elisabetta Cerbai e comprendente alcuni

docenti e ricercatori in rappresentanza delle diverse aree disciplinari, il Sistema Bibliotecario d'Ateneo (SBA), la FUP e il SIAF. La sistematicità e il rigore con i quali la commissione ha condotto la propria attività sotto la guida di Elisabetta Cerbai, decisamente coinvolta nella tematica dell'accesso aperto, hanno portato nel giro di pochi mesi a finalizzare alcuni obiettivi molto concreti. Il primo risultato in ordine temporale e fondamento di tutte le fasi successive è stata la definizione della "Policy di ateneo sull'open access"<sup>5</sup> che, dopo un ampio e approfondito confronto con la Commissione Ricerca, è stata approvata all'unanimità dal Senato Accademico nella seduta del 11 aprile 2012. Successivamente, grazie alla sensibilità del rettore Alberto Tesi, è avvenuto il recepimento dell'impegno a favore dell'accesso aperto nel nuovo Statuto dell'Ateneo, entrato in vigore l'8 maggio 2012, che al comma 2 dell'Art.8 "Ricerca scientifica" recita: «L'Università informa la disciplina delle attività di ricerca ai principi della trasparenza e della pubblicità; fa propri i principi dell'accesso pieno e aperto alla letteratura scientifica e promuove la libera diffusione in rete, nei circuiti della comunità scientifica internazionale, dei risultati delle ricerche prodotte in Ateneo». Infine, il lavoro della commissione si è focalizzato sulla individuazione delle soluzioni più idonee per dotare l'ateneo di un repository istituzionale destinato a divenire lo strumento per promuovere l'accesso aperto ai prodotti della ricerca e in grado di supportare adeguatamente i docenti ed i ricercatori. L'approccio seguito ha privilegiato l'aspetto funzionale rispetto a quello tecnico, ovvero invece di procedere direttamente alla scelta dello strumento informatico attraverso il quale implementare il repository si è data la priorità alla valutazione della situazione esistente e dei requisiti funzionali, evitando così di far prevalere le scelte relative alla piattaforma tecnologica sulle esigenze di carattere funzionale e operativo. L'analisi svolta ha posto in evidenza che

---

<sup>5</sup>[http://www.unifi.it/notiziario/upload/sub/2012\\_2/policy\\_open\\_access.pdf](http://www.unifi.it/notiziario/upload/sub/2012_2/policy_open_access.pdf).

il costituendo repository istituzionale avrebbe dovuto soddisfare i seguenti requisiti funzionali:

- a. *Integrare il flusso di lavoro dei docenti e ricercatori.* Le attività di alimentazione e gestione del repository nella modalità self-archiving devono essere ben integrate ed armonizzate con i workflow di lavoro abituali degli autori (docenti e ricercatori) senza aggiungere oneri operativi;
- b. *Evitare il doppio inserimento dei dati.* Nell'ambito del requisito precedente occorre prioritariamente escludere il rischio di inserimenti plurimi dei metadati e dei full text dei documenti;
- c. *Valorizzare U-GOV Catalogo Ricerca.* Per l'Ateneo U-GOV Catalogo Ricerca è uno strumento ormai consolidato e con un contenuto rilevante di prodotti/pubblicazioni registrate che rappresentano il risultato del lavoro della comunità scientifica dell'Ateneo e che comprendono numerosi full-text dichiarati ad accesso pubblico;
- d. *Facilitare la ricerca.* L'interfaccia di ricerca e consultazione deve essere semplice ed intuitiva e possibilmente allineata con alcuni servizi ben conosciuti (es. Scopus);
- e. *Assicurare la disseminazione dei metadati.* Occorre garantire l'integrazione con i metamotori di ricerca di tipo harvesting basati sul protocollo OAI-PMH;
- f. *Rendere omogeneo e coerente il contenuto.* Il repository deve contenere ed esporre le pubblicazioni relative all'attività di ricerca svolta dai docenti e ricercatori dell'Ateneo nei vari ambiti disciplinari.

In particolare l'approfondimento dei primi tre requisiti ha messo in evidenza il ruolo di U-GOV Catalogo Ricerca che è stato attivato

presso l'ateneo di Firenze da aprile 2008 e che nel periodo autunno 2011 – primavera 2012 è stato lo strumento usato per il conferimento, da parte di docenti e ricercatori, dei dati relativi alle proprie pubblicazioni (metadati descrittivi e full text) ai fini della Valutazione della Qualità della Ricerca (VQR) secondo le linee guida dettate dall'Agazia di Valutazione del Sistema Universitario e della Ricerca (ANVUR). Di conseguenza la modalità di inserimento in U-GOV Ricerca è divenuta in quel periodo familiare alla maggior parte degli autori e il catalogo non si è arricchito soltanto dei metadati, la cui registrazione è stata agevolata dalle possibilità di cattura tramite Digital Object Identifier (DOI) e di importazione da formati standard (e.g. BibText, Endnote, Institute for Scientific Information (ISI), Metadata Object Description Schema (MODS)) forniti dagli editori, ma anche del full text di molte pubblicazioni tra cui diverse rese disponibili ad accesso aperto immediatamente o in modalità differita, con la specifica di una data di fine embargo. Questa realtà ha portato a riflettere sui vantaggi e gli svantaggi di affiancare a U-GOV Ricerca un secondo strumento software per implementare il repository open access. Tra i vantaggi si è notata certamente una maggior flessibilità dal punto di vista della configurabilità dei metadati (per quanto anche U-GOV offra ampie possibilità di specificare attributi descrittivi per ogni tipologia di pubblicazione) secondo un approccio più orientato all'ambito bibliografico e tenuto conto che i software utilizzati per la gestione di repository istituzionali (es. eprints.org, DSpace, Fedora Commons) consentono l'adozione di set di metadati estesi rispetto al Dublin Core non qualificato. Un ulteriore vantaggio offerto da tali software sarebbe stata la possibilità di definire dei work-flow attraverso i quali sottoporre le pubblicazioni depositate tramite self-archiving a uno o più passi di verifica, revisione e integrazione dei metadati da parte di esperti prima di essere rese visibili pubblicamente. Questa modalità avrebbe consentito la

manutenzione qualitativa (correzioni, integrazioni, arricchimento) dei metadati bibliografici da parte di bibliotecari competenti nelle varie aree disciplinari. Tra gli svantaggi di un doppio contenitore è emersa la necessità di costringere gli autori a operare su due interfacce diverse (U-GOV Catalogo Ricerca e il repository) non solo per quanto riguarda il layout grafico, ma soprattutto per quanto concerne le modalità di interazione (sequenza di immissione, controlli e vincoli). D'altra parte, U-GOV consente di svolgere agevolmente le operazioni di self-archiving e, per gli atenei che come Firenze lo hanno adottato, è anche lo strumento primario per mantenere aggiornato il sito docente MIUR e pertanto il suo utilizzo costante è divenuto irrinunciabile ed è entrato a far parte del normale flusso di lavoro dei docenti e dei ricercatori. Un secondo svantaggio emerso nell'analisi dell'adozione del doppio repository è rappresentato dalla duplicazione dei metadati e dei full-text che prima di tutto implicano la maggiore occupazione di risorse hardware. Questa motivazione potrebbe essere ritenuta di scarsa rilevanza visto il costo limitato delle risorse hardware, tuttavia si devono considerare i maggiori costi operazionali Operating Expenditure (OpEx) dovuti al mantenimento e alla gestione di due diverse piattaforme tecnologiche. Inoltre occorre valutare anche l'aspetto funzionale, ovvero il fatto che un doppio contenitore/ambiente avrebbe comportato l'inevitabile disallineamento dei metadati per effetto degli interventi di manutenzione dei metadati da parte degli autori e dei team di manutenzione bibliografica sul secondo repository rispetto a quelli originari di U-GOV Ricerca che sono sincronizzati con la banca dati MIUR. Per quanto riguarda l'arricchimento e la manutenzione dei metadati, tenuto conto delle dimensioni dell'ateneo di Firenze (24 Dipartimenti e oltre 1900 tra docenti e ricercatori strutturati che operano su molti settori disciplinari) e delle scarse risorse umane disponibili, è stata valutata anche la difficoltà di attivare team di

bibliotecari in grado di assicurare con continuità la manutenzione qualitativa dei metadati e curarne costantemente l'arricchimento, previa ovviamente la preventiva interazione con gli autori delle pubblicazioni. In U-GOV l'autore che registra un prodotto è anche il "proprietario" e gestore dei metadati e comunque sono previste figure di utenti privilegiati in grado di intervenire su di essi. Infine, rispetto al requisito dell'omogeneità e della coerenza dei contenuti, la commissione si è orientata verso una specializzazione per archivi dedicati. Così il materiale inerente la didattica (dispense, slide, registrazioni audio e video) è stato fatto confluire sulla piattaforma E-learning, da tempo resa operativa da SIAF con il software open source Moodle. Altre tipologie di dati e documenti non strettamente connessi alla produzione scientifica sono ospitati nelle "Collezioni digitali" curate dallo SBA. FLORe si caratterizza dunque come "il Repository Istituzionale della Ricerca" con la doppia interpretazione dell'acronimo FLORe come "FLOResearch" e "FLORepository".

### **2.3 FLORe in sintesi**

Sulla base dei requisiti e delle considerazioni esposti nel paragrafo precedente per assolvere al proprio ruolo funzionale di repository istituzionale della produzione scientifica dell'Università di Firenze, FLORe è stato concepito come una estensione di U-GOV Catalogo Ricerca. In sostanza FLORe è a tutti gli effetti un Repository Open Access che, dal punto di vista logico, sviluppa pienamente il modello originario, già individuato nel 2000, fondato sulle due componenti, Data Provider e Service Provider, dell'architettura distribuita OAI. Il Data Provider, ovvero il nucleo centrale di FLORe, è U-GOV Catalogo Ricerca che fornisce agli autori l'interfaccia e le funzionalità per l'inserimento e la manutenzione dei metadati, per il caricamento dei full-text e per la loro caratterizzazione come documenti ad accesso aperto, con o senza embargo. In questo modo gli autori conservano

immutato il work-flow a cui sono abituati e viene loro evitata la necessità di utilizzare un secondo strumento per l'immissione dei dati bibliografici e delle pubblicazioni, inoltre si evita ogni duplicazione dei metadati e dei contenuti che sono mantenuti in un unico repository fisico. Infine il Data Provider conserva anche le funzionalità aggiuntive, rilevanti per i docenti e i ricercatori, che sono intrinseche di U-GOV Ricerca ovvero:

- i. evitare la digitazione manuale delle descrizioni bibliografiche importandole da formati standard o catturando i record bibliografici dai siti degli editori tramite DOI;
- ii. il riversamento automatico delle pubblicazioni immesse o modificate nella banca dati personale del docente/ricercatore grazie all'integrazione con il sito docente MIUR;
- iii. il "riconoscimento" dell'autore negli archivi delle risorse umane, operando l'aggancio con i dati di carriera (matricola, ruolo etc.) e garantendo così l'integrazione del catalogo ricerca con le informazioni del dominio del personale dell'Ateneo.

Il Data Provider è dotato dell'interfaccia OAI-PMH che Consorzio interuniversitario per il Calcolo Automatico (CINECA) ha implementato nel corso del 2011-2012 su richiesta di Firenze e di altri atenei come l'Università di Verona (Gabaldo). Tramite tale protocollo avviene l'harvesting dei metadati da parte di Pleiadi, il motore di ricerca nazionale per l'accesso centralizzato alla letteratura scientifica depositata negli archivi aperti italiani, istituiti e gestiti da università ed enti di ricerca, sviluppato dal Consorzio Interuniversitario Lombardo per L'Elaborazione Automatica (CILEA) e gestito da CINECA dopo l'unificazione dei consorzi universitari.<sup>6</sup> Infine Il

---

<sup>6</sup><http://www.openarchives.it/pleiadi>.

Service Provider è costituito dall'interfaccia di ricerca e visualizzazione dei risultati realizzata e mantenuta autonomamente da SIAF. Con il contributo dello SBA e il coordinamento di Mauro Guerrini, in base alle specifiche individuate e ai test effettuati nell'ambito del Sottogruppo Tecnico della Commissione Open Access, le funzioni di ricerca sono state sviluppate per versioni successive raffinando le funzionalità e perfezionando gli algoritmi a partire dalla versione Beta che è stata presentata ufficialmente il 25 ottobre 2012 in occasione della settimana dell'Accesso Aperto 2012. Attualmente la maschera di ricerca (Fig.1) propone quattro macro canali: "Titolo, Abstract, Keywords, Autore", "Titolo, Abstract, Keywords", "Autore", "Titolo". I primi due sono stati concepiti per fornire ricerche immediate "alla Google", ovvero sulle singole parole immesse: in tal modo è possibile indicare nel campo di ricerca parole che appartengono a domini diversi. Nei canali "Autore" e "Titolo" la ricerca avviene invece sulla frase esatta (per esempio cercando Mario Rossi in Autore si trovano solo le pubblicazioni in cui compare la stringa "Mario Rossi" nel campo "Autore"). È possibile utilizzare gli operatori logici AND, OR, NOT se indicati in una frase posta tra virgolette. Per raffinare la ricerca sono disponibili alcuni filtri:

- anno di pubblicazione: consente di selezionare un singolo anno oppure un intervallo di tempo da... a...);
- data inserimento: permette di visualizzare le pubblicazioni inserite in un determinato periodo (es. ultimi 15/30/90/180 gg, anno corrente, anno precedente);
- tipologia: il menù a tendina permette di scegliere il tipo di pubblicazione cui si è interessati (es. contributo su rivista, contributo in volume, tesi di dottorato etc.).

Il motore di ricerca prevede anche l'utilizzo di wild card con effetto di generalizzare la ricerca:

- “%” sostituisce uno o più caratteri;
- “\_” indica una posizione singola, all’interno di una parola, in cui può apparire un qualunque carattere;
- “,” consente di trovare i risultati, in base al canale prescelto, in cui è presente almeno una occorrenza di uno dei termini indicati. I risultati sono ordinati in base al numero di termini trovati;
- “\$” permette di cercare parole che hanno la stessa radice linguistica.

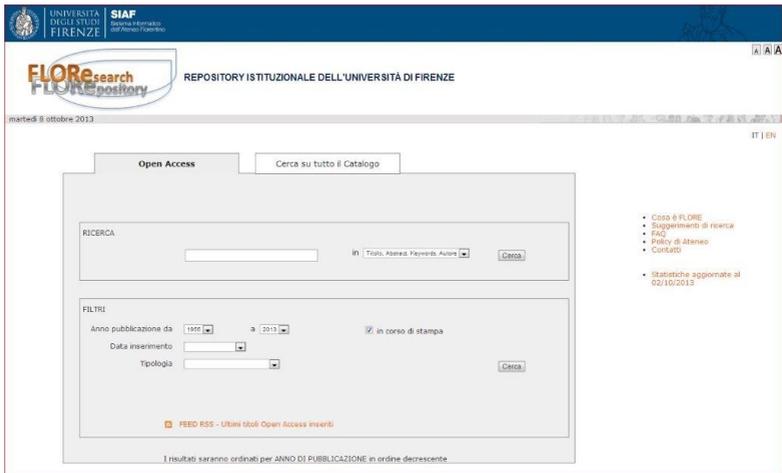


Figura 1: Maschera di ricerca

Se nella stringa di ricerca si specificano parole che non sono indicizzate (stop words, a seconda del canale prescelto e degli eventuali operatori booleani immessi, la ricerca può dare luogo a risultati

nulli. Al momento della stesura di questo articolo è in corso l'assemblamento della lista di queste parole e la valutazione se eliminarle completamente. I risultati della ricerca (Fig.2) vengono proposti ordinati per anno di pubblicazione decrescente, quindi mostrando prima le pubblicazioni più recenti, preceduti da icone grafiche che permettono di individuare a colpo d'occhio la tipologia della pubblicazione e sono presentati in pagine successive se il numero supera una soglia prefissata. La descrizione bibliografica sintetica (Fig.3) è cliccabile per consentire la consultazione della scheda bibliografica di dettaglio che propone tutti i campi descrittivi inseriti in U-GOV. Per il full text open access viene mostrato il formato (di solito PDF), la dimensione ed il link "Download" che permette di visualizzare/catturare il documento. Se disponibili i link al sito dell'editore e il DOI vengono sempre visualizzati in conformità a quanto previsto dalle politiche degli editori relative ai diritti di autore<sup>7</sup>. Infine, segue l'indicazione di quali autori sono "Autori dell'Università di Firenze" in base al loro precedente auto-riconoscimento in U-GOV. Da ciascuno di questi è possibile procedere con una ricerca implicita che recupera e propone tutte le pubblicazioni ad accesso aperto dell'autore appartenente all'Ateneo.

Con la stessa modalità ed interfaccia, ma con alcuni filtri aggiuntivi, cliccando sul tab "Cerca su tutto il Catalogo" è possibile effettuare ricerche su tutte le pubblicazioni, ovvero non più esclusivamente su quelle ad accesso aperto. Infine, un brevissimo cenno alla piattaforma tecnologica. I dati bibliografici utilizzati per le ricerche sono derivati da alcune tabelle virtuali (viste) esposte dal database U-GOV opportunamente materializzate e indicizzate con il motore Oracle Text. La soluzione di avvalersi di Oracle non deve destare sorpresa poiché tutti i database del sistema informativo dell'Ateneo, incluso U-GOV, usano il Data Base Management System (DBMS)

---

<sup>7</sup><http://www.sherpa.ac.uk/romeo>.

Oracle; quindi, per ragioni di uniformità ed efficienza oltre che per il know-how esistente, questa è stata la soluzione naturale. Le procedure di interrogazione sono scritte in Java (JSP e servlet) e gestite tramite la piattaforma open source standard Apache-Tomcat, ormai consolidata da oltre 10 anni per lo sviluppo di tutti i servizi online per studenti e docenti. L'accesso ai full-text avviene per mezzo di web-service appositamente esposti da U-GOV: in tal modo i contenuti (file PDF) vengono prelevati e trasferiti solo quando si effettua l'operazione di download.

### 3 Valutazioni, prospettive e conclusioni

Al 4 novembre 2013 FLORe comprende oltre 6.000 pubblicazioni con full-text open access, di cui 560 sono tesi di dottorato, rispetto al totale di poco più di 129.100 pubblicazioni/prodotti registrate/i in U-GOV Catalogo Ricerca. Dal punto di vista dei metadati, la principale criticità riscontrata è la notevole eterogeneità con la quale sono state registrate le stringhe autore nelle descrizioni bibliografiche. Solo in minima parte è stata usata la forma completa "Nome Cognome"; spesso sono state usate altre forme: iniziale del nome puntata, seguita dal cognome, con o senza uno spazio separatore, oppure il cognome seguito da uno spazio e dall'iniziale del nome puntato. Il separatore delle stringhe autore di solito è la virgola, ma anche per questo vi sono delle eccezioni. Questa varietà di forme può dare luogo a risultati incompleti nella tipica ricerca per autore formulata come "Nome Cognome" (es. Mario Rossi) ed è la conseguenza della difficoltà di assicurare nel tempo, fin dall'avvio del catalogo, uniformità per l'immissione della stringhe autore, tenuto conto del contesto self-archiving, dell'eterogeneità delle fonti bibliografiche e delle consuetudini vigenti nelle diverse aree disciplinari. Valutata l'impossibilità di intervenire ex-post per bonificare ed uni-

The screenshot displays the FLORe search interface. At the top, it identifies the repository as 'REPOSITORY ISTITUZIONALE DELL'UNIVERSITÀ DI FIRENZE'. The date is 'martedì 8 ottobre 2013'. The page is titled 'Ricerca - Risultati'. Under 'Pubblicazioni', it lists search criteria: 'contenenti guerrini data in Titolo, Abstract, Keywords, Autore', 'pubblicati dall'anno 1959 all'anno 2013 e in corso di Stampa', and 'con Full Text OPEN ACCESS'. It reports 'Trovati 11 pubblicazioni - ordinate per ANNO DI PUBBLICAZIONE in ordine decrescente' and 'Pubblicazioni da 1 a 11'. A 'LEGENDA' section defines icons for various document types. The main list shows the first three entries, each with a document icon, author names, and journal information.

Figura 2: Lista sintetica di risultati

The screenshot shows the 'Dettagli pubblicazione' page for a specific article. It includes the following details: 'Autore/i: Mauro Guerrini, Tatiana Possemato', 'Titolo: Linked data: un nuovo alfabeto del web semantico / Mauro Guerrini, Tatiana Possemato', 'Anno: 2012', 'Rivista: BIBLIOTECHE OGGI', 'Pagine: 7 - 15', 'ISSN: 0392-8586', and 'Tipologia: Contributo su rivista - Articolo su rivista'. The 'Keywords' are 'library linked data, web semantico'. The 'Abstract' discusses the Semantic Web and linked data. It provides a 'Full text depositato' link to a PDF file and a 'Permanent Link' to the document's URL. The author's name, MAURO GUERRINI, is highlighted in red.

Figura 3: Scheda bibliografica di una pubblicazione

formare le stringhe autore è stata considerata l'ipotesi di restringere la ricerca per autore ai soli autori "riconosciuti" (quelli per i quali, grazie all'aggancio con i dati delle risorse umane che fungono da authority file, sono disponibili il nome ed il cognome corretti), soluzione che è stata poi scartata in quanto non consentirebbe di cercare anche autori esterni all'Ateneo che in molte pubblicazioni di alcune discipline (es. quelle biomediche) sono assai numerosi. A questo limite pone in parte rimedio la funzione di ricerca per "autore riconosciuto" attivabile con un semplice click (link) sulla stringa "Nome Cognome" proposta nella lista "Autori dell'Università di Firenze" a piè di ciascuna scheda bibliografica di dettaglio. È incorso di valutazione anche la possibilità di arricchire la ricerca di FLORe con una funzionalità di browse per "autori riconosciuti" che consenta di recuperare immediatamente per ciascuno di essi le relative pubblicazioni ad accesso aperto. Rientra nelle prospettive a medio termine il processo di convergenza annunciato da CINECA tra U-GOV Catalogo Ricerca e Surplus, il software sviluppato e gestito da CILEA prima della confluenza dei consorzi universitari Consorzio interuniversitario per le Applicazioni di Supercalcolo Per Università e Ricerca (CASPUR) e CILEA in CINECA. Al momento della stesura di questo articolo le uniche informazioni disponibili sono quelle relative all'incontro CINECA Library User Group (CLUG) del 9 ottobre<sup>8</sup>. Per quanto riguarda FLORe non vi saranno impatti se, come dichiarato, CINECA manterrà invariata l'esposizione delle viste dei metadati e i web-services di U-GOV Ricerca. Eventuali modifiche architetturali conseguenti all'integrazione di DSpace in U-GOV (es. l'archiviazione dei full-text, tabelle dei metadati etc.) avranno certamente conseguenze a livello tecnico per gli atenei che come Firenze ospitano U-GOV sulle proprie piattaforme, ma non saranno perce-

---

<sup>8</sup>[https://clug.cineca.it/system/files/accesso/riunioni/Webex%20riunione%20Gruppo%20gestione%20dati%20finale\\_0.doc](https://clug.cineca.it/system/files/accesso/riunioni/Webex%20riunione%20Gruppo%20gestione%20dati%20finale_0.doc)

pite dagli utenti finali di FLORe. Diversamente, dal punto di vista funzionale ed operativo, nel caso in cui la convergenza con Surplus determinasse un cambiamento significativo nell'attuale interfaccia grafica di U-GOV Catalogo Ricerca, che è parte integrante di FLORe e della sequenza operativa ormai consolidata per l'inserimento di una pubblicazione e del relativo full-text da parte degli autori dell'ateneo fiorentino, l'impatto sui medesimi autori, docenti e ricercatori, che archiviano le loro pubblicazioni in FLORe sarebbe significativo. In questo caso, quindi, le evoluzioni e le conseguenze dovranno essere attentamente valutate con CINECA. Per quanto riguarda i contenuti del repository, la loro crescita è legata anche alle iniziative che saranno intraprese per promuovere il deposito ad accesso aperto da parte dei docenti e dei ricercatori dell'Ateneo, al fine di assecondare un sempre più ampio spostamento verso l'archiviazione delle pubblicazioni in modalità open access, anche tramite servizi di aiuto e consulenza per la cura dei metadati bibliografici e la verifica delle politiche degli editori. Infine, pur senza voler anticipare soluzioni da verificare e approfondire con più attenzione in futuro, tenuto conto del recente orientamento internazionale e governativo verso la pubblicazione di dati aperti (open data), FLORe è già predisposto anche per consentire la pubblicazione e la diffusione di set di dati relativi a una ricerca scientifica; i relativi file possono infatti essere allegati alla pubblicazione ed essere messi a disposizione in modalità open access secondo formati aperti. Nel contesto dei requisiti individuati e delle prospettive tratteggiate, l'implementazione del repository istituzionale FLORe è una soluzione efficace sia sul piano tecnico che su quello funzionale poiché estende e integra U-GOV Catalogo Ricerca minimizzando l'entità dei moduli applicativi appositamente sviluppati, dunque riducendo il costo gestionale, inoltre non altera le modalità di lavoro da parte dei docenti e dei ricercatori, evitando di distoglierli con altre interfacce e percorsi operativi dal loro work-

flow abituale per la descrizione e l'archiviazione delle pubblicazioni. Al tempo stesso favorisce la disseminazione dei metadati grazie al protocollo OAI-PMH e fornisce un'interfaccia di ricerca semplice ed intuitiva che consente la rapida consultazione dei contenuti.

## Riferimenti bibliografici

- Cotoneschi, Patrizia. «Firenze University Press fra distribuzione tradizionale e libero accesso». *L'archivio E-Prints dell'università di Firenze: prospettive locali e nazionali. Atti del Convegno*. A cura di Patrizia Cotoneschi. Firenze University Press, 2004. 13–22. (Cit. a p. 154).
- Gabaldo, Maria. «L'Open Archive e il suo sviluppo nell'Università di Verona». *Il Documento Digitale I/MMXIII*. (2013): 38–39. (Cit. a p. 161).
- Maraviglia, M. Giulia. «L'Archivio E-Prints nella Biblioteca digitale dell'Università di Firenze». *L'archivio E-Prints dell'università di Firenze: prospettive locali e nazionali. Atti del Convegno*. A cura di Patrizia Cotoneschi. Firenze University Press, 2004. 5–12. (Cit. a p. 154).
- Pasqui, Valdo. «Archivi di documenti elettronici: un modello di riferimento per la realizzazione della biblioteca ibrida». *Biblioteche Oggi* 18.10. (2000): 6–12. (Cit. a p. 152).
- Pepou, Giancarlo. «Introduzione». *L'archivio E-Prints dell'università di Firenze: prospettive locali e nazionali. Atti del Convegno*. A cura di Patrizia Cotoneschi. Firenze University Press, 2004. 1–3. (Cit. a p. 154).

GAIA INNOCENTI, SIAF–Sistema Informatico dell'Ateneo Fiorentino–Università di Firenze.

[gaia.innocenti@unifi.it](mailto:gaia.innocenti@unifi.it)

CRISTINA MUGNAI, SIAF–Sistema Informatico dell'Ateneo Fiorentino–Università di Firenze.

[cristina.mugnai@unifi.it](mailto:cristina.mugnai@unifi.it)

VALDO PASQUI, SIAF–Sistema Informatico dell'Ateneo Fiorentino–Università di Firenze.

[valdo.pasqui@unifi.it](mailto:valdo.pasqui@unifi.it)

---

Innocenti, G., C. Mugnai. V., V. Pasqui. "FLORe – Florence Repository. L'Archivio ad accesso aperto dell'Università di Firenze". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January

2014): Art: #9102. DOI: [10.4403/jlis.it-9102](https://doi.org/10.4403/jlis.it-9102). Web.

**ABSTRACT:** The University of Florence has a long and consolidated experience in the exploitation of open archives and self archiving by researchers. After a short introduction on current open access context in Europe and in Italy, the article recalls past experiences by tracing the history of the first e-prints archive and then describes the genesis of FLORe, the new research institutional repository developed to promote open access archiving by teachers and researchers at the University of Florence. Afterwards the architecture and the main functional and technical features of FLORe are discussed aiming to depict and to motivate the approach followed in implementing FLORe as an extension of U-GOV Research, the catalogue of scientific products, rather than an autonomous repository. Following and implementing the requirements stated by the University of Florence Open Access Commission, FLORe proves to be the right solution because it reduces operational costs, does not add a new graphical user interface, preserves the already well consolidated end-users self-archiving workflow and, at the same time, provides a simple and intuitive search interface with OAI-PMH support for research metadata dissemination.

**KEYWORDS:** E-research; Institutional repository; Metadata; Open access; Open archive; Research data.

---

Submission: 2013-10-08  
Accettazione: 2013-11-18  
Pubblicazione: 2014-01-01







# La ricerca bibliografica in infermieristica: un'esperienza di condivisione tra strutture della Fondazione IRCCS Policlinico S. Matteo di Pavia

Michele Chieppi, Piera Bergomi, Donatella De Amici

## 1 Introduzione

La complessità della medicina moderna e la rapidità con cui essa si evolve, impone che l'approfondimento continuo in ogni singola specialità diventi indispensabile (Perino e Ruffino). Le informazioni bio-mediche sono fondamentali in quanto sono la base su cui si fonda la ricerca e il sapere scientifico è strettamente legato alla formazione dei professionisti della salute (Singh e Garg). Tra questi professionisti risultano esserci gli infermieri, per i quali, la formazione rivolta alla ricerca permette loro di assimilare e creare nuove conoscenze scientifiche (National Research Council (US) Committee to Study the National Needs for Biomedical e Personnel). In tale contesto gli infermieri hanno avviato da tempo un percorso di approfondimento legato alla professionalizzazione dal punto di vista formativo, legislativo e di pratica clinica assumendo responsabilità



nell'area organizzativo gestionale e didattico-formativa; un tragitto che li vede coinvolti immancabilmente anche nella ricerca scientifica (Silvestro). La stessa Federazione Nazionale dei Collegi Infermieri (IPASVI) ha messo a disposizione di tutti gli infermieri italiani iscritti agli albi l'accesso gratuito al database Cinahl Plus with Full Text, strumento fondamentale per la formazione e aggiornamento in campo infermieristico. Uno degli obiettivi da perseguire nell'ambito dell'infermieristica italiana è infatti promuovere la pubblicazione di articoli scientifici (Dall'Oglio et al.).

## **2 L'informazione scientifica: la rivincita degli open-access**

Fonti di informazione sono disponibili in abbondanza on-line anche gratuitamente (Halijwale, Manjunath e Pujar) e per l'infermieristica internet è la porta verso un mondo di risorse apparentemente inesauribili (Jange e Lalitha) ma è importante saper distinguere la documentazione valida ed accreditata da quella che non presenta sufficienti caratteristiche da poter essere presa in considerazione in un contesto scientifico, come ad esempio i blog o siti non supportati da fondazioni o associazioni di settore conosciute. Ciò non implica ad esempio che siano da preferire le riviste internazionali accessibili unicamente attraverso la stipula di onerosi abbonamenti e corredate da un alto valore di Impact Factor (IF) a scapito di quelle free-access che, seppur corredate di IF e da un'indiscussa validità intellettuale, vengono spesso ed ingiustamente considerate come fonti da porsi in secondo piano. Questo discutibile modo di valutare la letteratura scientifica nasce da un percorso che trova le sue radici verso la metà del secolo scorso e per la precisione nel 1960 quando Eugene Garfield a Philadelphia fonda l'Institute for Scientific Information (ISI).

Le strategie di marketing da parte dell'ISI fecero credere che la presenza di una rivista nel suo elenco fosse sinonimo di scientificità e che l'IF fosse un criterio esaustivo per stabilirne sia la popolarità che il valore. Da qui, riviste consultate unicamente da specialisti del settore in causa, diventarono opere irrinunciabili per ogni biblioteca universitaria. Ovviamente, inteso questo, l'editoria scientifica alzò i prezzi senza nessun controllo (Pievatolo). Se il 1992 è la data dell'acquisto dell'ISI da parte della Thompson Scientific & Healthcare, gli anni '90 sanciscono una crescita considerevole degli autori che volevano pubblicare su riviste accompagnate da alti valori di IF che, di conseguenza, crebbero d'importanza. Gli autori sanno che è necessario pubblicare su riviste internazionali con un alto valore di IF, certificato da Thomson Reuters e farsi citare o pubblicare nel contesto ove figurano persone note, oppure su riviste molto citate con il fine di ottenere un alto punteggio sulla produttività scientifica individuale (h-index) e un altrettanto alto valore dell'impatto del singolo sulla comunità scientifica (g-index). Il panorama che ne consegue vede l'autore, in certi casi, indotto addirittura a pagare per poter pubblicare e parallelamente si vede una crescita economica considerevole delle multinazionali dell'editoria (Penco). Nel caso in cui l'autore (e questo è un fattore che si presenta nella maggior parte dei casi) sia anche un ricercatore, deve i suoi profitti non alla vendita dell'articolo ma all'impatto sulla comunità dei ricercatori, ossia al fatto di essere letto, utilizzato e citato e da terzi (De Robbio). Il percorso che porta ad una *rivincita* degli open-access, è lungo e tortuoso fino a culminare con due episodi significativi: il primo del maggio 2012 in cui Winsotn Hide, direttore associato di *Genomics*, si dimette e scrive una lettera al *Guardian* sottolineando il suo dissenso rivolto verso un sistema che offre alti profitti all'editore e nega l'accesso ai risultati ai ricercatori dei paesi in via di sviluppo; il secondo in cui Bart Knots direttore di *Acta Tropica* imita Winsotn Hide abbandonando il comita-

to di direzione. Le sue parole sono chiare: « I giorni dell'iperprofitto per la pubblicazione scientifica potrebbero essere contati» (Penco). In sintesi: tutti gli ostacoli ad un libero accesso ai contenuti sono una barriera che limita l'impatto scientifico nel contesto della comunità e conducono ad un rallentamento nella crescita della ricerca e al progresso economico e sociale in quanto vanno a ledere il processo formativo dell'individuo (De Robbio). La situazione per l'Infermieristica italiana si specchia in questa realtà: insieme alla necessità che le riviste si arricchiscano di articoli scientifici, è conveniente per gli autori infermieri pubblicare su riviste di ampia diffusione in modo da raggiungere un più vasto pubblico del settore e non solo la comunità scientifica (Dall'Oglio et al.). Similmente ad altre figure legate alla sanità infatti, anche quella dell'infermiere è una professione polivalente, la quale, può esprimersi in più settori tra cui quello della ricerca scientifica (Hardicre). L'editoria on-line ha cambiato il paesaggio attraverso il quale filtra la trasmissione dell'informazione accademica e la comunità infermieristica ha solo cominciando da poco ad accettare e utilizzare l'open-access prendendo coscienza che queste risorse facilitano l'apprendimento attivo e costante (Nick). Un aspetto non meno importante è in questo contesto il saper utilizzare correttamente gli strumenti informatici: («Impact of information technology (COMPUTERS) on biomedical information centres and libraries (ICLs) in India: a critical evaluation») affermano che l'introduzione di servizi basati sull'informatizzazione cambierà il futuro dei *biomedical information centres* e delle biblioteche. Siamo d'accordo con questa affermazione e crediamo ci sia bisogno di una figura professionale di riferimento che insegni come utilizzare al meglio il computer e come cercare correttamente on-line le risorse. Tuttavia, la ricerca professionale all'interno delle banche dati bio-mediche ed infermieristiche può risultare complessa; da qui, la necessità di servizi bibliotecari di qualità diventa pronunciata (Hallyburton e

St. John). In più, la fusione delle informazioni e la gestione delle competenze conoscitive del bibliotecario esperto in sanità unitamente all'esperienza clinica dell'infermiere può essere un grande valore per la comunità scientifica e valorizza la figura dell'infermiere nella ricerca (Chieppi e Bergomi).

### **3 Il progetto formativo della Fondazione IRCCS Policlinico San Matteo di Pavia**

Con il fine di condividere e approfondire la conoscenza delle strategie e delle tecniche di ricerca bibliografica, di trovare e valutare le informazioni provenienti da banche dati bibliografiche selezionate e siti internet accreditati, stimolare la capacità di compilazione della citazione delle fonti bibliografiche e favorire l'implementazione delle conoscenze acquisite a terzi (ad esempio permettere ai tutor-infermieri di insegnare agli studenti in infermieristica i principi fondamentali da osservare per impostare un'efficace ricerca bibliografica), è nato, in seno alla Fondazione Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS) Policlinico San Matteo, il progetto formativo: *La ricerca bibliografica in ambito infermieristico*. Tale intenzione si è sviluppata con la cooperazione fra Struttura Formazione e Sviluppo (SFS), la Biblioteca di Infermieristica (BIB), il Corso di Laurea (CdL) in infermieristica e il Servizio Infermieristico Tecnico e Riabilitativo Aziendale (SITRA) e si è rivolto ai tutor del CdL e ai componenti del SITRA. L'approccio collaborativo tra strutture, alla cui base esiste un progetto formativo, permettere al personale di un ente, istituzione, fondazione ecc. di imparare gli uni dagli altri e condividere le migliori pratiche. Permette inoltre alle strutture stesse e ai vari dipartimenti di soddisfare le esigenze di sviluppo della forza lavoro impegnata, promuovendo l'eccellenza clinica (Reay e Sears).

Condividendo la proposta di («Nursing research and bibliographic citation models») apparsa sulla *Revista Latino-Americana de Enfermagem* nel giugno 2009, secondo cui le riviste infermieristiche dovrebbero adottare un proprio modello di citazione a prescindere da come i professionisti pubblicano in medicina, un altro obiettivo del progetto è stato quello di stilare un manuale di regole (Chieppi), caratterizzato da un proprio stile bibliografico adatto alla citazione dei documenti maggiormente utilizzati nelle scienze infermieristiche (Chieppi, Bergomi e Mangiacavalli). Per raggiungere questi obiettivi è venuta in supporto la tipologia formativa di *Formazione sul Campo (FsC)* considerata in letteratura la forma più efficace di formazione, sia per le caratteristiche dell'*apprendimento dell'adulto*, sia per la metodologia utilizzata che evita la contrapposizione di ruoli tra insegnanti e discenti, sia per la necessaria definizione dei bisogni educativi che è essenziale alla sua stessa natura. La formazione è uno sviluppo sistematico delle conoscenze, delle abilità e degli atteggiamenti necessari ad un'organizzazione per raggiungere i suoi obiettivi (Talab e Tajafari): la fondazione San Matteo, infatti, in qualità di IRCCS, ha come *mission* quella di sostenere le iniziative formative a sostegno della ricerca scientifica al fine di garantire standard di qualità elevati ed uniformi. Il coordinamento, promosso anche dal settore formazione del sistema Bibliosan (Le Biblioteche in rete degli Enti Biomedici Italiani), è stato affidato alla Biblioteca di Infermieristica della Fondazione.

## 4 Materiali e metodi

È molto importante per i professionisti delle discipline bibliotecomiche soddisfare le aspettative degli utenti finali. Per conseguire questo obiettivo alcune qualità di base devono essere sviluppate: la chiarezza di pensiero, la buona relazione interpersonale e l'onestà

sono considerate più importanti di tutte le altre qualità (Regha e Rani). Abbiamo verificato che questo rapporto bibliotecario/utente può essere applicato anche nel binomio bibliotecario/tutor nel campo della condivisione delle conoscenze e abbiamo sostenuto l'idea che la sperimentazione e l'innovazione sono le fonti principali per creare nuovi saperi. Se ne evince che le organizzazioni devono offrire maggiore spazio agli studiosi e ai docenti per sperimentare nuovi progetti (Thakur e Thakur). Inoltre, un altro fattore che può essere considerato un nuovo punto di dialogo fra bibliotecario e tutor infermieristici, o meglio tra biblioteconomia e scienze infermieristiche, sono le cinque leggi formulate da Ranganathan. Considerando il fatto che le leggi stanno trovando applicazione al di là del confine della biblioteconomia, dall'area religiosa a quella ospedaliera (Sen), suggeriamo un modello per l'infermieristica:

1. l'assistenza infermieristica è per i pazienti;
2. ad ogni infermiere il suo paziente;
3. ad ogni paziente il suo infermiere;
4. migliorare la vita del paziente;
5. un ospedale cresce con gli infermieri per i pazienti.

L'arco temporale in cui si è sviluppato il progetto è stato di due anni, dall'inizio del 2011 alla fine del 2012. E' stato suddiviso in due livelli: uno base [Liv A] attivato nel 2011 ed uno avanzato di approfondimento [Liv B] nel 2012. L'accesso al Liv B è stato riservato a chi ha superato con profitto il Liv A. Entrambi i livelli hanno previsto un ciclo di 4 riunioni della durata di almeno 2 ore con una verifica finale dell'apprendimento. Il Liv A è stato ripetuto per 4 edizioni, il Liv B per 2. Le riunioni si sono svolte presso l'aula informatica della SFS dove i partecipanti hanno avuto individualmente a disposizione un

personal computer. Come tecnica didattica è stata utilizzata la lezione frontale, per diffondere le conoscenze iniziali, seguita in larga parte dal lavoro individuale, anche al di fuori delle riunioni e da esecuzioni dirette da parte dei partecipanti. La gestione operativa del corso è stata affidata al responsabile della BIB e la responsabilità scientifica al coordinatore del CdL.

*Obiettivo di apprendimento del Liv A:* Diffondere e condividere un linguaggio comune per la ricerca bibliografica: conoscere le principali strategie di ricerca e strumenti (banche dati, cataloghi nazionali e internazionali, motori di ricerca a carattere scientifico ecc.); arrivare alla condivisione di un protocollo relativo alle norme citazionali con la conseguente realizzazione di un manuale di regole e stile. Si riporta qui di seguito nel dettaglio la descrizione di ogni incontro con il fine di poter permettere di adattare e applicare la formula di condivisione dei saperi ad altre realtà e ad altre discipline. Si specifica che la stretta osservanza del programma di ogni riunione è stato mantenuto invariato in modo da poter stilare con la massima precisione i risultati e le statistiche finali. Le 4 riunioni delle 4 edizioni del Liv A hanno scrupolosamente rispettato il seguente ordine di condivisione dei saperi:

*Conoscenze e abilità tecnico-specialistiche condivise*  
*I riunione—insegnamenti ed attività svolte:*

1. Cenni di principi di bibliografia e biblioteconomia:

- definizione di bibliografia e biblioteconomia;
- tipologie delle bibliografie: aperte, chiuse, le nazionali, bibliografie di bibliografie;
- definizione di catalogo;
- tipologie di cataloghi: per autore, per soggetto ecc.;

- confronto tra i concetti di bibliografia e catalogo: similitudini e differenze;
  - spiegazione delle funzioni degli opac e dei meta-opac.
2. Esempi guidati di strategie di ricerca semplice ed avanzata in opac e meta-opac: opac d'ateneo dell'Università degli Studi di Pavia, opac Sbn, Mai Azalai, KvK Karlsruhe Virtual Catalog:
    - esempi frontali;
    - ricerca individuale secondo le strategie indicate;
    - ricerca all'interno delle singole riviste in lingua italiana.
  3. Discussione sulla formula da utilizzare per una corretta citazione delle fonti in ambito delle Scienze Infermieristiche.
  4. La citazione bibliografica: analisi delle strutture e formule per la stesura (monografie in bibliografia e in nota, articoli in bibliografia e in nota).

*I Riunione—descrizione dell'incontro:*

Introdotte le note operative relative alla FsC, si sono illustrati gli obiettivi da perseguire:

1. la creazione di un protocollo relativo alle regole e allo stile di una citazione bibliografica unica per i tutor-infermieri, i componenti del SITRA, gli infermieri della fondazione;
2. lo studio di strategie di ricerca all'interno degli opac, meta-opac, dei siti di pertinenza infermieristica e delle principali riviste infermieristiche disponibili tramite abbonamento online.

Al punto 1, la lezione ha avuto carattere frontale; dal punto 2 i partecipanti sono stati invitati a condividere in maniera attiva la lezione

operando individualmente e applicando le conoscenze impartite. I criteri stabiliti per eseguire la ricerca on-line dei documenti sono stati rispettati all'unanimità. Ha ottenuto particolare consenso dei partecipanti la ricerca delle fonti attraverso l'utilizzo del Karlsruhe Virtual Catalog (KVK). Sono stati segnalati i siti internet di carattere infermieristico più significativi e più affidabili dal punto di vista scientifico dei contenuti e le riviste consultabili e scaricabili in formato elettronico in lingua italiana. E' stata assegnata a fine riunione un'esercitazione con il fine di verificare il livello di apprendimento. Essa consisteva nella composizione di alcune citazioni bibliografiche (in nota e in bibliografia) partendo da dati identificativi espressi in ordine sparso.

*Il Riunione—insegnamenti ed attività svolte:*

1. Individuazione e consultazione dei seguenti siti internet di stimata credibilità scientifica per la consultazione e l'aggiornamento:
  - Infermieri.net, federazione IPASVI, Associazione nazionale infermieri di area critica (Aniarti), Associazione Nazionale Infermieri per la Prevenzione delle Infezioni Ospedaliere (Anipio), Nursearea, Associazione Italiana Infermieri di area Oncologica (AIIO), Clinica Virtuale, Centro per la valutazione dell'efficacia dell'assistenza sanitaria (CeVeAs), Gruppo Italiano per la Medicina Basata sulle Evidenze (GIMBE), Biblioteca Medica Virtuale, Scirus, Accademia di Scienze Infermieristiche, Rete Infermieri Oncologia, altre risorse bibliografiche in rete.
2. Individuazione e consultazione dei seguenti siti internet di stimata credibilità scientifica quali fonti di documentazione:

- Evidencebasednursing.it, Trip Database, banche dati di linee guida (Scottish Intercollegiate Guidelines Network (SIGN), New Zealand Guidelines Group, Institute for Clinical Systems Improvement, Agree, CeVeAs, Sistema Nazionale Linee Guida (SNLG) ecc.), banche dati di Revisioni sistematiche (Centre for Reviews and Dissemination (CRD) database, Bandolier, Joanna Briggs Institute ecc.), Istituti per l'Evidence-Based Nursing (EBN) e l'Evidence-Based Medicine (EBM) (GIMBE, Best Evidence Topics, The National Institute for Clinical Excellence ecc.), riviste on-line (On-line Journal of Rural Nursing and Health Care, Free Medical Journals ecc.).
3. Dal quesito alla strategia: gli operatori booleani:
    - gli operatori booleani AND, OR, NOT.
  4. Gli stili citazionali più utilizzati dalle riviste scientifiche internazionali:
    - Vancouver Style, Apa Style 6th Ed., Chicago Citation Style 16th Ed., Harvard Style in rapporto allo stile citazionale in fase di studio.
  5. Presentazione banche dati primarie e secondarie:
    - Medline, PubMed, Cinahl, The Cochrane Library;
    - principali funzioni comuni: limiti, ricerca avanzata ecc.
  6. La ricerca in PubMed: ricerca semplice e applicazione delle principali funzioni:
    - limits, advanced search, entrez: the Life Sciences Search Engine, Search in Books, Images, NCBI Web Site ecc.

7. Simulazioni guidate di ricerca in PubMed.
8. La ricerca in Cinahl: ricerca semplice e applicazione delle principali funzioni:
  - ricerca nel titolo, abstract, all-text, perfezionamento dei risultati, cattura pdf degli articoli, citazioni ecc.
9. Simulazioni guidate di ricerca in Cinahl.

*II Riunione—descrizione dell'incontro:*

Risolti i dubbi inerenti l'appuntamento precedente, l'incontro ha avuto carattere frontale dal punto 1 al punto 5 descritti in: *Insegnamenti ed attività svolte*. I partecipanti hanno poi interagito dal punto 6 al punto 9 in più simulazioni guidate in cui hanno partecipato attivamente all'individuazione di articoli in osservanza dei criteri di ricerca precedentemente stabiliti.

*III Riunione— insegnamenti ed attività svolte:*

1. Griglie per la valutazione critica degli studi (studi di trattamento, degli studi eziologici, degli studi prognostici, degli studi diagnostici, delle revisioni), calcolatori (per studi randomizzati e controllati, per studi di coorte, per studi caso-controllo, per studi diagnostici, per valori di AGREE standardizzati), virtual books;
2. Analisi, implementazione di strategie di ricerca nelle riviste elettroniche internazionali disaggregate per editore ed esercitazione guidata (Elsevier-Science Direct, Sage, Journal of the American Medical Association (JAMA), The New England Journal of Medicine, British Medical Journal, Blackwell-Wiley, BioMedCentral);

3. Analisi, implementazione di strategie di ricerca nel sistema di supporto alle decisioni cliniche (Best Practice del British Medical Journal (BMJ) Evidence Center);
4. Il Medical Subject Headings (MeSH) di PubMed (spiegazione della funzionalità e utilizzo del vocabolario controllato per l'indicizzazione degli articoli delle riviste biomediche);
5. La ricerca nella Cochrane Library;
6. La ricerca in Medline: ricerca semplice e applicazione delle principali funzioni (ricerca avanzata, lingua di pubblicazione, modalità di ricerca, limitazione dei risultati, cartella delle preferenze, perfezionamento dei risultati, source types, age, subject: major heading, gender, publication, ricerca visiva: group result, sort result, filter results by date, display style, fogli informativi basati sulle prove di efficacia, intestazioni ecc.);
7. La ricerca individuale in PubMed, Cinahl, Medline, Cochrane Library (individuare e descrivere la strategia di ricerca corretta, catturare la notizia bibliografica ed esprimerla correttamente, descrivere e motivare i criteri di inclusione ed esclusione dei documenti consultati).

### *III Riunione—descrizione dell'incontro:*

Per tutta la sua durata, la riunione ha coinvolto attivamente i partecipanti, i quali hanno interagito in più simulazioni guidate. Sono stati analizzati i siti di pertinenza infermieristica che a livello mondiale elaborano documenti di accreditato valore scientifico, puntando l'attenzione sulle riviste elettroniche e sugli editori internazionali. Stimolati dal desiderio di approfondire quanto appreso, i partecipanti si sono cimentati in ricerche personalizzate, usufruendo delle banche dati biomediche ed infermieristiche suggerite.

*IV Riunione—insegnamenti ed attività svolte:*

I partecipanti hanno sostenuto la prova finale secondo i parametri qui di seguito descritti:

- carattere individuale;
- ogni partecipante ha dovuto stabilire un argomento di ricerca nell'ambito delle scienze infermieristiche;
- la ricerca è stata effettuata all'interno delle banche dati analizzate durante gli incontri, avendo cura di riservare la priorità a Cinahl, PubMed, Cochrane Library, Medline;
- sono stati ricercati almeno 3 articoli. Questi, sono poi stati analizzati in modo da poterli includere od escludere dalla propria indagine (ha fatto fede la preventiva enunciazione del soggetto della ricerca ad inizio incontro);
- sono stati descritti nel dettaglio tutti i passaggi effettuati per il raggiungimento di ogni singola notizia bibliografica ed i sistemi utilizzati (advanced search, limits, operatori booleani, scelta del tipo di materiale ecc.). In aggiunta i partecipanti hanno dovuto riportare la striscia di ricerca (es./ nel caso di PubMed la search details) onde permettere di replicare la strategia;
- per ogni notizia bibliografica ritenuta valida, i partecipanti hanno elaborato la corretta citazione secondo la regola studiata nelle precedenti riunioni;
- ogni partecipante ha riprodotto in forma manoscritta quanto riportato nei punti precedenti.

Alla fine della prova, i partecipanti sono stati invitati a compilare liberamente un documento cartaceo in cui esprimevano il proprio e

personale parere sul ciclo delle riunioni. Si è pensato di far pervenire il documento in forma anonima con il fine di non influenzare il libero pensiero del singolo. Nella fase finale della riunione si è proceduti alla discussione finale stilando un bilancio riassuntivo.

*IV Riunione—descrizione dell'incontro:*

Dagli scritti consegnati in forma anonima, sommati alla discussione finale, i partecipanti hanno manifestato la necessità di nuovi percorsi di ricerca bibliografica da effettuarsi periodicamente, con il fine di approfondire e mantenere vive le competenze raggiunte. Il grado di soddisfazione appare collocato ad un buon livello, mentre pareri positivi si sono riscontrati a proposito dell'utilità del progetto e sulla sua gestione. Assodate le necessità sopra descritte riguardo a questa parte del percorso e alla condivisione di strategie comuni di ricerca, i presenti si sono dimostrati favorevoli all'implementazione di un protocollo condiviso che determini regole precise alla stesura della citazione bibliografica.

*Obiettivo di apprendimento del Liv B:* Approfondire lo studio sulla lettura e sulla qualità dell'articolo scientifico e condividere, attraverso esercitazioni pratiche, le conoscenze e le abilità da trasmettere a terzi (colleghi, infermieri, studenti ecc.) elaborando più percorsi di ricerca bibliografica su temi infermieristici concordati. In più, da un'analisi delle criticità effettuata all'interno del CdL, dopo esser stata pianificata una strategia di ricerca comune tra i tutor nelle precedenti edizioni del progetto di formazione, si necessitava di approfondimenti, aggiornamenti sul campo e di esercitazioni, con il fine di formare e abilitare i partecipanti a condividere quanto appreso con gli studenti del CdL. Si riporta qui di seguito nel dettaglio la descrizione di ogni incontro con il fine di poter permettere di adattare e applicare la formula di condivisione dei saperi ad altre realtà e ad altre discipline. Si specifica che la stretta osservanza del

programma di ogni riunione è stato mantenuto invariato in modo da poter stilare con la massima precisione i risultati e le statistiche finali. Le 4 riunioni delle 2 edizioni del Liv B hanno scrupolosamente rispettato il seguente ordine di condivisione dei saperi.

*Conoscenze e abilità tecnico-specialistiche condivise:*

*I Riunioni—insegnamenti ed attività svolte:*

1. Il MeSH di PubMed:

- definizione e descrizione del MeSH;
- la struttura gerarchica del MeSH (MeSH Tree);
- proprietà e funzioni;
- impostazione teorica delle strategie di ricerca;
- i subheading dei termini MeSH;
- porte d'accesso al MeSH: MeSH database, more resources, home National Centre for Biotechnology Information (NCBI);
- le funzioni del do not include MeSH terms;
- il major topic;
- la combinazione del MeSH con l'advanced search e con i limits.

2. Esempi guidati di strategie di ricerca utilizzando il MeSH di PubMed:

- esempi frontali;
- ricerca individuale secondo le strategie indicate.

3. Condivisione di due differenti percorsi di ricerca da affrontarsi nel corso della corrente edizione.

4. Percorso guidato sul I caso di strategia condiviso:

- tecniche di distrazione per la gestione del dolore in età pediatrica;
- traduzione e parole chiave;
- ricerca delle evidenze scientifiche.

5. Individuazione delle revisioni sistematiche, lettura degli abstract, lettura parziale delle parti dell'articolo che più interessano la strategia.

6. Condivisione degli obiettivi.

*I Riunione—descrizione dell'incontro:*

Introdotte le note operative relative alla FsC, si sono illustrati gli obiettivi da perseguire:

1. approfondire le principali strategie di ricerca già rese note (banche dati, cataloghi nazionali e internazionali, motori di ricerca a carattere scientifico);
2. esercitarsi sull'elaborazione di una ricerca bibliografica su temi infermieristici concordati.

Ad inizio riunione si è trattato degli argomenti già affrontati nella precedente edizione: dai caratteri identificativi delle banche dati utilizzate (PubMed, Cinahl, Cochrane ecc.), all'applicazione delle funzioni di limits e advanced search ecc. Al punto 1, descritto nell'elenco degli *Insegnamenti ed attività svolte*, la lezione ha avuto carattere frontale: si sono resi noti gli approfondimenti teorici. Dal punto 2 i partecipanti sono stati invitati a condividere in maniera attiva la lezione operando individualmente e applicando le strategie secondo criteri guidati. Particolare interesse ha suscitato la scelta del percorso di ricerca stabilito: Tecniche di distrazione per la gestione

del dolore in età pediatrica. La strategia di ricerca ha mirato ad individuare in primo luogo delle revisioni sistematiche che riportassero evidenze scientifiche riguardo la sperimentazione di tecniche di distrazione non farmacologiche nell'area sopraindicata. Da qui l'individuazione di altre revisioni che le presentassero quantitativamente. Valutata l'autorità scientifica del materiale trovato, si è risaliti all'individuazione degli articoli primari che supportassero le evidenze per ogni singola tecnica di distrazione individuata. Nel dettaglio il percorso ha tenuto conto dell'importanza dello stabilire un quesito di ricerca iniziale e delle relative parole chiave, concentrandosi anche sulla compatibilità delle medesime rispetto al tesaurus (dizionario controllato di termini) MeSH. I partecipanti hanno poi condiviso la lettura di tutti gli abstract presentati, di parti di articoli primari e di revisioni valutate positivamente per essere incluse nel percorso di ricerca. L'avvenuta inclusione/esclusione dei documenti si è svolta valutando i parametri relativi alle opzioni offerte dalla funzione limits ((ages, dates, type of article, languages, menu delle search field tags) alla pertinenza delle keywords–abstract–title. Gli approfondimenti appresi attraverso la pratica del MeSH sono stati poi messi in relazione e fatti interagire nella medesima strategia di ricerca con le funzioni di advanced search e limits, comunque rese già note (ma oggi aggiornate, riviste e approfondite) nell'edizione precedente.

*Il Riunione—insegnamenti ed attività svolte:*

Sono stati analizzati e discussi i seguenti documenti, che andiamo a citare seguendo la regola prodotta nella prima edizione del percorso di formazione:

- CHAMBERS C.T., TADDIO A., UMAN L.S., McMURTRY C.M., HELPinKIDS TEAM, "Psychological interventions for reducing pain and distress during routine childhood immu-

nizations: a systematic review”, *Clin Ther.*, 2009;31 Suppl 2:S77–S103.

- COHEN L.L., BLOUNT R.L., COHEN R.J., SCHAEN E.R., ZAFF J., “Comparative study of distraction versus topical anesthesia for pediatric pain management during immunizations”, *Health Psychol.*, 1999;18:591–598.
- FOWLER–KERRY S., LANDER J.R., “Management of injection pain in children”, *Pain*, 1987;30:169–175.
- GONZALEZ J.C., ROUTH D.K., ARMSTRONG F.D., “Effects of maternal distraction versus reassurance on children’s reactions to injections”, *J Pediatr Psychol.*, 1993;18:593–604.
- HARRISON D., YAMADA J., ADAMS–WEBBER T., OHLS–SON A., BEYENE J., STEVENS B., “Sweet tasting solutions for reduction of needle–related procedural pain in children aged one to 16 years”, *Cochrane Database Syst Revl.*, 2011 Oct 5;(10):CD008408.
- HOFFMAN H.G., PATTERSON D.R., MAGULA J., CARROUGHER G.J., ZELTZER K., DAGADAKIS S., SHARAR S.R., “Water–friendly virtual reality pain control during wound care”, *J Clin Psychol.*, 2004;60:189–195.
- KUTTNER L., “Management of young children’s acute pain and anxiety during invasive medical procedures”, *Pediatrician*, 1989;16(1–2):39–44.
- MANIMALA M., BLOUNT R.L., COHEN L.L., “The influence of parental reassurance and distraction on children’s reactions to an aversive medical procedure”, *Child Health Care*, 2000;29 :161–177.

- MASON S., JOHNSON M.H., WOOLLEY C., "A comparison of distractors for controlling distress in young children during medical procedures", *J Clin Psychol Med Settings*, 1999;6:239–248.
- MURPHY G., "Distraction techniques for venepuncture: a review", *Paediatr Nurs.*, 2009 Apr;21(3):18–20.
- PRINGLE B., HILLEY L., GELFAND K., DAHLQUIST L.M., SWITKIN M., DIVER T., SULC W., ESKENAZI A., "Decreasing child distress during needle sticks and maintaining treatment gains over time", *J Clin Psychol Med Settings*, 2001;8:119–130.
- SCHECHTER N.L., ZEMPSKY W.T., COHEN L.L., McGRATH P.J., McMURTRY C.M., BRIGHT N.S., "Pain reduction during pediatric immunizations: evidence-based review and recommendations", *Pediatrics*, 2007 May;119(5):e1184–98.
- SHARAR S.R., MILLER W., TEELEY A., SOLTANI M., HOFFMAN H.G., JENSEN M.P., PATTERSON D.R., "Applications of virtual reality for pain management in burn-injured patients", *Expert Rev Neurother.*, 2008 Nov;8(11):1667–74.
- SPARKS L., "Taking the ouch out of injections for children: using distraction to decrease pain", *Am J Matern Child Nurs.*, 2001;26:72–78.
- UMAN L.S., CHAMBERS C.T., McGRATH P.J., KISLEY S., "Psychological interventions for needle-related procedural pain and distress in children and adolescents", *Cochrane Database Syst Rev.*, 2006 Oct 18;(4):CD005179.
- VESSEY J., CARLSON K., MCGILL J., "Use of distraction with children during an acute pain experience", *Nurs Res.*, 1994;43:369–372.

- YASTER M., “Multimodal analgesia in children”, *Eur J Anaesthesiol.*, 2010 Oct;27(10):851–7.

*II Riunione—descrizione dell’incontro:*

In riferimento al quesito: Tecniche di distrazione per la gestione del dolore in età pediatrica, e in relazione al percorso effettuato nella precedente riunione, si è provata l’esistenza di evidenze scientifiche che supportino l’utilizzo delle tecniche di distrazione correlate alla gestione del dolore nei bambini di età prescolare (compresi nella fascia da 2 a 5 anni secondo i parametri pubblicati da PubMed nell’area limits–ages). Tale conoscenza è stata consolidata attraverso la lettura degli abstract, di parti dei full–text e discussione in aula. Da una seconda revisione sistematica ottenuta attraverso un nuovo percorso di ricerca, che ha enumerato le tecniche di distrazione utilizzate, sono stati individuati 9 articoli che supportano con evidenze scientifiche ciascuna delle distrazioni. Ad ogni partecipante è stato consegnato il compito di esaminarlo a fondo. La tesi che ogni partecipante dovrà sostenere nel prossimo incontro è l’aggiornamento scientifico relativo ad ogni articolo, vale a dire cercare, citare per esteso (secondo gli insegnamenti impartiti nel primo ciclo di incontri) ed esporne i contenuti in relazione alla documentazione scientifica compresa tra la data di pubblicazione dell’articolo in oggetto e il febbraio 2012. E’ stato poi condiviso da tutti il simulare altri percorsi di ricerca con il fine di consolidare gli insegnamenti impartiti.

*III Riunione—insegnamenti ed attività svolte:*

- Presentazione, condivisione e discussione degli elaborati;
- Implementazione di un nuovo percorso di ricerca.

*III Riunione—descrizione dell’incontro:* In riferimento al percorso effettuato nella precedente riunione, i partecipanti hanno sviluppato e approfondito quanto richiesto. Durante la lezione, i presenti hanno

presentato e discusso il proprio lavoro. Si è quindi in sintesi arrivati, grazie all'apporto del lavoro di tutti, ad un aggiornamento relativo alle tecniche di distrazione non farmacologiche relazionate al dolore nel bambino in età prescolare. Si è proseguito risolvendo dubbi ed incertezze relative ai percorsi studiati. Archiviato il primo percorso, si è proseguito implementandone un secondo: Le infezioni correlate del catetere venoso centrale nell'adulto e nel bambino. In questa parte della riunione, i partecipanti hanno condiviso una comune strategia di ricerca e l'hanno applicata.

#### *IV Riunione—insegnamenti ed attività svolte:*

I partecipanti hanno sostenuto la prova finale secondo i parametri qui di seguito descritti:

- carattere individuale;
- ogni partecipante ha dovuto stabilire un argomento di ricerca nell'ambito delle scienze infermieristiche;
- la ricerca è stata effettuata all'interno delle banche dati analizzate durante gli incontri, avendo cura di riservare la priorità a Cinahl, PubMed, Cochrane Library, Medline;
- sono stati ricercati almeno 3 articoli. Questi, sono poi stati analizzati in modo da poterli includere od escludere dalla propria indagine (ha fatto fede la preventiva enunciazione del soggetto della ricerca ad inizio incontro);
- sono stati descritti nel dettaglio tutti i passaggi effettuati per il raggiungimento di ogni singola notizia bibliografica ed i sistemi utilizzati (advanced search, limits, operatori booleani, scelta del tipo di materiale ecc.). In aggiunta i partecipanti hanno dovuto riportare la striscia di ricerca (es./ nel caso

di PubMed la search details) onde permettere di replicare la strategia;

- per ogni notizia bibliografica ritenuta valida, i partecipanti hanno elaborato la corretta citazione secondo lo stile studiato nella precedente edizione;
- ogni partecipante ha riprodotto in forma manoscritta quanto riportato nei punti precedenti.

#### *IV Riunione—descrizione dell'incontro:*

Il grado di soddisfazione appare collocato ad un buon livello, mentre pareri positivi si sono riscontrati a proposito dell'utilità del progetto e sulla sua gestione. I partecipanti si sono auto-valutati in grado di implementare ad altri le conoscenze apprese.

#### *Risultati*

*Liv A:* Dei 31 iscritti iniziali, 20 (64.5%) hanno completato l'intero percorso partecipando alla prova finale. Di questi, il 100% ha ricevuto 8 crediti Educazione Continua in Medicina (ECM).

All'unanimità, l'implementazione del nuovo stile bibliografico ha ottenuto il consenso dei partecipanti. Alcune difficoltà sono da riscontrarsi nella stesura del campo identificativo della pubblicazione e dell'ordine di stesura. In generale è stato possibile stimare che le competenze raggiunte, già apprezzabili, erano sicuramente migliorabili con il contributo di un protocollo avente carattere manualistico con formule ed esempi. Il manuale contenente le nuove regole è stato approvato e condiviso da tutti i partecipanti e dai responsabili delle strutture afferenti. Si è quindi ritenuto utile metterlo a disposizione di tutti gli infermieri della fondazione e di tutte le figure gravitanti intorno al CdL di Pavia. Al termine del I° incontro di ogni edizione, i partecipanti sono stati invitati a compilare un questionario a risposta multipla in forma anonima con 5 quesiti che

<b>Liv A</b>	<b>1. Ed.</b>	<b>2. Ed.</b>	<b>3. Ed.</b>	<b>4. Ed.</b>	<b>Totali</b>
<b>Iscritti</b>	<b>10</b>	<b>8</b>	<b>7</b>	<b>6</b>	<b>31</b>
I Riunione	8	7	4	5	24
II Riunione	7	7	4	4	22
III Riunione	7	7	4	5	23
IV Riunione	7	6	4	4	21
<b>Amm. prova finale</b>	<b>7</b>	<b>6</b>	<b>4</b>	<b>3</b>	<b>20</b>
<b>Accreditati</b>	<b>7</b>	<b>6</b>	<b>4</b>	<b>3</b>	<b>20</b>

**Tabella 1:** Iscritti al Liv A, dettaglio presenze, ammessi alla prova finale, accreditati con ECM

indagavano le conoscenze pregresse e 4 che miravano a verificare il livello di apprendimento dopo l'ascolto della riunione. I partecipanti che hanno risposto sono in totale 23. Dai risultati, coloro che hanno avuto modo di citare una fonte bibliografica sono il 100%, di cui il 78.3% autogestendosi, il rimanente 21.7% interpellando un esperto del settore. Sull'esistenza dei cataloghi elettronici (opac, meta-opac), il 47.8% ne è a conoscenza, di questi li hanno utilizzati il 13%, non ne ha mai sentito parlare il 52.2%. Sull'utilizzo delle banche dati mediche ed infermieristiche risponde positivamente il 52.2%, non vengono utilizzate dal 13%, vengono consultate raramente dal 34.8%. La documentazione proposta dai siti on-line di pertinenza infermieristica è presa in considerazione dal 91.3% contro un rimanente 8.7% che li consulta raramente. Nella seconda parte del questionario ha riconosciuto la citazione corretta l'82.6%, la definizione di bibliografia e quella di catalogo è chiara al 91.3%. Riconosce l'esempio di letteratura grigia il 95.7%.

Le prova finale del Liv A ha risposto ai seguenti criteri: osservare il carattere di lavoro individuale, produrre un quesito iniziale

<b>Questionario: Domande/Risposte</b>	<b>Si</b>	<b>No</b>	<b>Raramente o aiutati</b>
Avete mai scritto una citazione bibliografica?	18	0	5
Conoscete i cataloghi elettronici?	11	12	0
Se sì, li avevate mai utilizzati?	3	20	0
Utilizzate le banche dati?	12	13	8
Utilizzate siti infermieri- stici?	21	0	2

**Tabella 2:** Prima parte del questionario riservato alla verifica delle conoscenze pregresse.

<b>Questionario: Domande/Risposte</b>	<b>Giusto</b>	<b>Sbagliato</b>
<b>Individuare la citazione corretta</b>	19	14
<b>Definizione di bibliografia</b>	21	2
<b>Definizione di catalogo</b>	21	2
<b>Riconoscere l'esempio di lett. grigia</b>	22	1

**Tabella 3:** Seconda parte del questionario riservato alla verifica di quanto appreso.

di ricerca, utilizzare le banche dati analizzate durante gli incontri, applicare i criteri di inclusione/esclusione sui risultati ottenuti, descrivere nel dettaglio tutti i passaggi effettuati per il raggiungimento di ogni singola notizia bibliografica e le opzioni di ricerca utilizzate, appuntare i passaggi per permettere di replicare la strategia in fase di correzione, elaborare la citazione per ogni notizia bibliografica

ritenuta utile, riprodurre in forma manoscritta quanto riportato nei punti precedenti.

## 4.1 La regola

Validata nel 2013 e pubblicata nelle sue linee essenziali (Chieppi, Bergomi e Mangiacavalli) da *L'Infermiere: Notiziario Aggiornamenti Professionali*, organo ufficiale della IPASVI che l'ha rese visibili in open-access, nasce espressamente per soddisfare più esigenze. Innanzitutto in letteratura è possibile apprezzare stili di citazione bibliografica studiati appositamente per alcune categorie professionali e per l'infermieristica non è mai stata proposta una regola precisa. Ne è un esempio l'APA Style, elaborato in seno ad una precisa associazione di categoria, l'American Psychological Association e le cui norme di pubblicazione sono da tale associazione sostenute. Un altro tipo di esigenza nasce dall'analisi relativa al fermento, cambiamenti di stile e personalizzazioni significative legate all'area delle citazioni bibliografiche visibili nelle riviste internazionali di infermieristica. Se in passato il Vancouver style e l'Harvard style erano gli stili più comuni utilizzati per le citazioni nelle riviste infermieristiche internazionali, c'è oggi una tendenza crescente a preferire l'Apa Style seguito dal Vancouver style. Vi sono anche le esigenze editoriali della singola rivista a determinare situazioni ibride apportando modifiche più o meno significative a stili bibliografici statuari (Chieppi, Bergomi e Mangiacavalli) ad entrare di forza in questo contesto. Nel successivo elenco proposto dall'ISI, aggiornato a dicembre 2012, si riportano le riviste in ordine decrescente per valore di IF alle prime 36 posizioni a fianco delle quali si è trascritta la segnatura dello stile bibliografico utilizzato, avendo cura di segnalare con la scritta *Modello personalizzato* le riviste in cui abbiamo individuato l'utilizzo di un proprio stile indipendente.

	<b>Rivista</b>	<b>IF</b>	<b>Stile bibliografico</b>
1	ONCOL NURS FORUM	2.509	APA 6 <sup>th</sup> Ed.
2	BIRTH-ISS PERINAT C	2.182	Vancouver
3	INT J NURS STUD	2.178	Modello personalizzato
4	CANCER NURS	1.792	AMA 10 <sup>th</sup> Ed.
5	MIDWIFERY	1.777	Modello personalizzato
6	EUR J CARDIOVASC NUR	1.711	Vancouver
7	RES NURS HEALTH	1.708	APA 6 <sup>th</sup> Ed.
8	J PEDIATR HEALTH CAR	1.661	APA 6 <sup>th</sup> Ed.
9	AM J CRIT CARE	1.656	AMA 10 <sup>th</sup> Ed.
10	NURS OUTLOOK	1.522	APA 6 <sup>th</sup> Ed.
11	J NURS SCHOLARSHIP	1.490	APA 6 <sup>th</sup> Ed.
12	J ADV NURS	1.477	Modello personalizzato
13	ADV SKIN WOUND CARE	1.438	AMA 10 <sup>th</sup> Ed.
14	J CARDIOVASC NURS	1.431	AMA 10 <sup>th</sup> Ed.
15	J NURS ADMIN	1.419	Vancouver
16	EUR J ONCOL NURS	1.410	Modello personalizzato
17	NURS RES	1.402	APA 5 <sup>th</sup> Ed.
18	J PERINAT NEONAT NUR	1.364	AMA 9 <sup>th</sup> Ed.
19	HEART LUNG	1.318	AMA 10 <sup>th</sup> Ed.
20	PERSPECT PSYCHIATR C	1.298	APA 6 <sup>th</sup> Ed.
21	BIOL RES NURS	1.278	APA 6 <sup>th</sup> Ed.
22	NURS EDUC TODAY	1.241	Modello personalizzato
23	WORLDV EVID-BASED NURS	1.239	APA 5 <sup>th</sup> Ed.
24	APPL NURS RES	1.219	APA 5 <sup>th</sup> Ed.
25	J NURS CARE QUAL	1.193	AMA 10 <sup>th</sup> Ed.
26	WESTERN J NURS RES	1.189	APA 6 <sup>th</sup> Ed.
27	J NURS MANAG	1.181	Modello personalizzato
28	J MIDWIFERY WOM HEAL	1.163	AMA 10 <sup>th</sup> Ed.
29	J HUM LACT	1.150	AMA 10 <sup>th</sup> Ed.
30	PAIN MANAG NURS	1.145	APA 5 <sup>th</sup> Ed.
31	J WOUND OSTOMY CONT	1.142	AMA 10 <sup>th</sup> Ed.
32	MCN-AM J MATERN-CHIL	1.123	APA 6 <sup>th</sup> Ed.
33	AM J NURS	1.119	APA 6 <sup>th</sup> Ed.
34	J CLIN NURS	1.118	Modello personalizzato
35	J ASSOC NURSE AIDS C	1.090	APA 6 <sup>th</sup> Ed.
36	NURS CRIT CARE	1.082	Modello personalizzato

**Figura 1:** Elenco delle prime 36 riviste infermieristiche in ordine decrescente per valore di IF. A fianco la descrizione dello stile bibliografico utilizzato (dicembre 2012).

La regola è stata applicata ad un'ampia tipologia di documenti ed enumera un totale di 48 casi analizzati: dalla semplice monografia con un solo autore al volume di un'opera in più tomi con titolo diverso da quello dell'opera generale; dalle citazioni tratte da enciclopedie e dizionari alle citazioni di tutto il materiale relativo alla legislazione italiana. Il tutto è stato poi compreso in un manuale redatto in formato elettronico e messo a disposizione degli infermieri, autori, studenti ecc. afferenti alla Fondazione IRCCS Policlinico S. Matteo di Pavia. Per ogni caso preso in considerazione si è fornita spiegazione della formula applicata con relativi esempi distinguendo la citazione in bibliografia generale da quella a piè di pagina o nel corpo del testo (la regola prevede infatti entrambe le figure). Si vedano ad esempio le tabelle rappresentate nelle figure 2 e 3 nella pagina successiva.

Informazioni sul documento	Norme stilistiche	Esempio
COGNOME Nome dell'Autore,	COGNOME Nome,	MARTUCCIELLO G.,
"Titolo dell'Articolo : ed eventuale sottotitolo",	"Stampatello minuscolo posto tra virgolette : eventuale sottotitolo",	"Hirschsprung's disease, one of the most difficult diagnoses in pediatric surgery: a review of the problems from clinical practice to the bench",
<i>Titolo del periodico : ed eventuale sottotitolo,</i>	<i>Corsivo : eventuale sottotitolo in corsivo,</i>	<i>Eur J Pediatr Surg,</i>
Anno di edizione e mese (solo se esplicitamente indicato);	anno Mese;	2008;
Numero dell'annata o del volume	Usare solo l'espressione numerica. Es./ 3	18
Numero del fascicolo, posto fra parentesi rotonda, seguito da due punti. Es./ (...):	Usare solo l'espressione numerica. Es./ (5):	(3):
Indicazione dei numeri di pagina dell'Articolo all'interno del Periodico.	Usare solo l'espressione numerica. Es./ 22-34.	55-59.

**In Bibliografia:**

MARTUCCIELLO G., "Hirschsprung's disease, one of the most difficult diagnoses in pediatric surgery: a review of the problems from clinical practice to the bench", *Eur J Pediatr Surg*, 2008;18(3):55-59.

**Figura 2:** Articolo di Periodico, in Bibliografia

Nel caso in cui la nota sia da porsi nel corpo del testo: (COGNOME e iniziale del nome: numero di pagina o di pagine)

Informazioni sul documento	Norme stilistiche	Esempio
Iniziale del Nome e COGNOME,	N. COGNOME,	G. MARTUCCIELLO,
“Titolo dell’ Articolo [senza indicare il sottotitolo]”,	“Stampatello minuscolo posto tra virgolette”,	“Hirschsprung’s disease, one of the most difficult diagnoses in pediatric surgery”,
<i>Titolo del Periodico</i> [senza indicare il sottotitolo],	<i>Corsivo</i> ,	<i>Eur J Pediatr Surg</i> ,
Anno di edizione e mese (solo se esplicitamente indicato);	anno Mese;	2008;
Numero dell’annata o del volume	Usare solo l’espressione numerica. Es./ 3	18
Numero del fascicolo, posto fra parentesi rotonda, seguito da due punti. Es./ (...):	Usare solo l’espressione numerica. Es./ (5):	(3):
Indicazione dei numeri di pagina dell’Articolo all’interno del Periodico.	Usare solo l’espressione numerica. Es./ 22-34.	55-59.

**In Nota:**  
G. MARTUCCIELLO, “Hirschsprung’s disease, one of the most difficult diagnoses in pediatric surgery”, *Eur J Pediatr Surg*, 2008;18(3):55-59.

**Figura 3:** Articolo di periodico, in nota

*Nel corpo del testo:*

(MARTUCCIELLO G.:55–59)

Si è pensato, a differenza di molti stili noti in letteratura, di sostituire (nella seconda parte della citazione da porsi nel corpo del testo) la data di pubblicazione con il numero o i numeri di pagina corrispondenti in quanto si è appurato che nel caso di ripetizione della citazione della medesima monografia all’interno, ad esempio, di una tesi di laurea, è fondamentale per il relatore individuare la parte precisa del volume a cui si è fatto riferimento. Nel caso invece in cui la composizione debba essere compresa in un articolo di una rivista che adotti la nostra regola, il sistema risulterebbe ugualmente funzionale. Molto spesso infatti nelle scienze infermieristiche, nel momento in cui si tratta di uno specifico argomento, vengono citati documenti redatti dal medesimo autore; per questo se dovesse coincidere sia il cognome che la data di composizione bisognerebbe apporre le lettere dell’alfabeto onde distinguere le citazioni in causa, andando anche a modificare nella sua statuaria stesura la citazione

posta in bibliografia generale. La nostra variante permette di mantenere inalterata la fisionomia della citazione in bibliografia generale e distinguerebbe con la stessa efficacia i documenti citati nel corpo del testo. Un'altra evidente particolarità è rappresentata dal fatto che il cognome dell'autore o degli autori sono posti in stampatello maiuscolo: questo per rendere più visibile la fonte all'interno del testo e per donare maggiore dinamicità grafica alla composizione, in particolar modo, se si tratta di una tesi di laurea generalmente stampata con i tipi di carattere più comuni dei personal computer sui tradizionali fogli in formato A4. L'indice informatizzato del manuale permette infine al lettore di individuare il preciso punto all'interno del nostro lavoro in cui viene riportato il tipo di documento che lo interessa. Dall'analisi dei risultati presentati, giunge ancora più significativa la necessità di proporre per le scienze infermieristiche una propria regola di citazione con il fine di distinguere la letteratura di quest'area da quella relativa alle altre discipline comprese nell'insieme delle scienze legate alla sanità (Chieppi, Bergomi e Mangiacavalli).

*Liv B:* Dei 17 iscritti iniziali, 14 (77.7%) hanno completato l'intero percorso partecipando alla prova finale. Di questi, il 100% ha ricevuto 8 crediti ECM.

Fra le strategie proposte, ci si è concentrati particolarmente sul percorso di ricerca inerente alle *Tecniche di distrazione non farmacologiche per la gestione del dolore in età pediatrica*. Nel dettaglio, sia questo percorso che gli altri affrontati, hanno tenuto conto dell'importanza dello stabilire un quesito di ricerca iniziale e delle relative parole chiave concentrandosi anche sulla compatibilità delle medesime rispetto al *Thesaurus MeSH*. Identificata una review di *Pediatrics* del 2007 (Schechter et al.) che ne elenca un discreto numero (film, bolle di sapone, caleidoscopio, brevi fiabe, musica ecc.) si è chiesto ad ognuno dei partecipanti di sceglierne una ed in base alla citazione

<b>Liv B</b>	<b>1. Ed.</b>	<b>2. Ed.</b>	<b>Totali</b>
<b>Iscritti iniziali</b>	<b>9</b>	<b>8</b>	<b>17</b>
I Riunione	9	7	16
II Riunione	9	6	15
III Riunione	9	6	15
IV Riunione	9	5	14
<b>Amm. prova finale</b>	<b>9</b>	<b>5</b>	<b>14</b>
<b>Accreditati</b>	<b>9</b>	<b>5</b>	<b>14</b>

**Tabella 4:** Iscritti al Liv B, dettaglio presenze, ammessi alla prova finale, accreditati con ECM.

bibliografica che riportava all'articolo originale, il quale sosteneva con evidenze scientifiche la tecnica, di aggiornare lo stato dell'arte. I partecipanti hanno poi condiviso la lettura degli articoli ricercati e selezionati. La prova finale del Liv B, ha osservato i medesimi criteri utilizzati precedentemente nella prova del Liv A.

## 5 Discussione

Per quanto riguarda il Liv A, dai dati del questionario somministrato al termine di ogni edizione e dalla customer, è risultato che i partecipanti necessitavano di una conoscenza e di un'analisi dettagliata dei cataloghi elettronici on-line e le loro funzioni. Nel contempo si dimostravano ampiamente soddisfatti e chiedevano espressamente un secondo livello con approfondimenti. Il manuale di citazione bibliografica prodotto è stato approvato e condiviso dai partecipanti e dalle strutture coinvolte. Nel progetto di Liv B, i partecipanti, grazie anche all'esperienza maturata nel livello precedente, hanno saputo approfondire le proprie conoscenze relative alle opzioni di ricerca

avanzata, raffinato la tecnica di ricerca e maturato ulteriormente la capacità di valutare quantitativamente e qualitativamente la letteratura scientifica. I risultati finali pertanto sono stati tutti molto positivi. Il successo del percorso formativo è legato a diversi fattori: prima di tutto all'arco temporale in cui si è svolto, rappresentando la programmazione della formazione del personale che, uscendo da ogni tentazione di occasionalità, si è proiettata sul medio-lungo periodo. In secondo luogo il successo è legato alla tipologia formativa utilizzata. Così come definito dalla D.g.r. 12 aprile 2006–n. 8/2372 la formazione *on the job* (equivalente di FsC) rappresenta una tipologia di formazione professionale che permette l'interazione tra apprendimento riflessivo sulle pratiche finalizzato al loro rinnovamento attraverso nuove conoscenze e competenze esperte e l'apprendimento trasformativo della professionalità degli operatori per corrispondere a cambiamenti interni all'organizzazione sanitaria e al miglioramento dello stato di salute della popolazione di riferimento. La FsC favorisce l'apprendimento degli adulti, in quanto pratica formativa capace di accrescere la motivazione ad apprendere coinvolgendo e responsabilizzando direttamente i partecipanti.

## 6 Conclusioni

La professione infermieristica ha compiuto negli ultimi anni moltissimi progressi sia in ambito scientifico che accademico. La spinta al sapere ed a tesoriare i saperi deve sempre essere vista come un momento topico e determinante nello stimolare la capacità di condividere e quindi conferire circolarità ai saperi stessi. Questo progetto formativo ne rappresenta un esempio, dove diverse strutture hanno lavorato, condiviso e raggiunto obiettivi considerevoli valorizzando il lavoro dei colleghi. Sapere pertanto dove sia la scienza, equivale ad avere le soluzioni per risolvere gli innumerevoli problemi che

tutte le figure professionali nel campo dell'assistenza si trovano a fronteggiare quotidianamente nella pratica clinica, nella gestione dei servizi o nella didattica. L'infermiere di oggi deve essere sempre più competente nell'esercizio della professione e deve conoscere quali interventi abbiano superato le prove di efficacia del metodo scientifico e quali invece abbiano mostrato risultati dubbi o addirittura nulli o dannosi. Anche se oggi la divulgazione scientifica è molto diffusa e alla portata di tutti, le conoscenze in ambito sanitario permangono molto complesse. I due livelli del progetto hanno portato a condividere la conoscenza avanzata di tutte le funzioni e le opzioni offerte delle banche dati mediche ed infermieristiche primarie e secondarie e di sviluppare la capacità di implementare strategie di ricerca complesse. Congiuntamente, i partecipanti sono preparati a compilare notizie bibliografiche citandole secondo le regole dettate dai principali stili bibliografici internazionali. Il contributo di tutti i partecipanti ha reso possibile la realizzazione di un manuale di regole. Il progetto ha quindi risposto fattivamente alla necessità di pianificare una strategia di ricerca comune con il fine di diffondere e condividere un linguaggio collettivo per la ricerca bibliografica infermieristica in contesti educativi, professionali e gestionali in diverse strutture afferenti la Fondazione IRCCS Policlinico San Matteo di Pavia.

## Riferimenti bibliografici

- Angordans, Jordi Piqué, Ramon Camaño Puig e Carmen Piqué Noguera. «Nursing research and bibliographic citation models». *Revista Latino-Americana de Enfermagem* 17.3. DOI: [10.1590/S0104-11692009000300017](https://doi.org/10.1590/S0104-11692009000300017). (2009): 387–395. (Cit. a p. 178).
- Chieppi, Michele. *Regole per la redazione della citazione bibliografica*. Fondazione IRCCS Policlinico S.Matteo di Pavia, 2011. (Cit. a p. 178).
- Chieppi, Michele e Piera Bergomi. «Biblioteche di infermieristica: proposta di un dialogo multidisciplinare». *Biblioteche Oggi* 31.4. (2013): 50–55. (Cit. a p. 177).

- Chieppi, Michele, Piera Bergomi e Barbara Mangiacavalli. «La citazione bibliografica nella letteratura infermieristica: una nuova Regola dalla Biblioteca di infermieristica della Fondazione Ircs Policlinico S. Matteo di Pavia». *L'infermiere: Notiziario Aggiornamenti Professionali, Organo ufficiale della Federazione Nazionale dei Collegi Ipasvi* 2. (2013): 13–19. (Cit. alle pp. 178, 198, 202).
- Dall'Oglio, Immacolata, et al. «Lettura delle riviste infermieristiche e partecipazione all'attività scientifica da parte degli infermieri: risultati di uno studio osservazionale». *L'Infermiere: Notiziario Aggiornamenti Professionali, Organo ufficiale della Federazione Nazionale dei Collegi Ipasvi* 47.2. (2010): 25–33. (Cit. alle pp. 174, 176).
- De Robbio, Antonella. *Archivi aperti e comunicazione scientifica*. A cura di Clio Press. Napoli, 2007. (Cit. alle pp. 175, 176).
- Halijwale, Sageta S, G.K. Manjunath e S.M. Pujar. «Electronic journals: modalities for providing free access». *Annals of Library and Information Studies* 51.2. (2004): 82–85. (Cit. a p. 174).
- Hallyburton, Ann e Barbara St. John. «Partnering with your library to strengthen nursing research». *J Nurs Educ* 49.3. DOI: [10.3928/01484834-20091118-04](https://doi.org/10.3928/01484834-20091118-04). (2010): 164–167. (Cit. a p. 176).
- Hardicre, Jayne. «An exploration of the role of the research nurse and its impact». *British Journal of Nursing* 22.3. (2013): 168–169. (Cit. a p. 176).
- Jange, Suresh e K. Sami Lalitha. «Influence of Internet on library and information centres of National Institutes of Technology in India». *Annals of Library and Information Studies* 53.4. (2006): 184–197. (Cit. a p. 174).
- National Research Council (US) Committee to Study the National Needs for Biomedical, Behavioral e Clinical Research Personnel. *Research Training in the Biomedical, Behavioral, and Clinical Research Sciences*. A cura di National Academies Press (US). Washington (DC), 2011. [http://www.nap.edu/openbook.php?record\\_id=12983](http://www.nap.edu/openbook.php?record_id=12983). (Cit. a p. 173).
- Nick, Jan. «Open Access Part II: The Structure, Resources, and Implications for Nurses». *Online Journal Of Issues In Nursing* 17.1. DOI: [10.3912/OJIN.Vol17No01PPT031](https://doi.org/10.3912/OJIN.Vol17No01PPT031). (2012). (Cit. a p. 176).
- Penco, Carlo. «Scribi al servizio delle multinazionali». *Il Sole 24 Ore*. (2012). <<http://faberblog.ilssole24ore.com/2012/01/scribi-al-servizio-delle-multinazionali-della-ricerca/>>. (Cit. alle pp. 175, 176).
- Perino, Bruno e Emanuele Davide Ruffino. «La relazione medico-paziente: riflessione su un rapporto in crisi». *Torino Medica* 1. (2012): 12–21. (Cit. a p. 173).
- Pievatolo, Maria Chiara. «Riviste all'Indice. La controriforma dell'Anvur». *Roars: Return On Accademic Research*. (2012). <<http://www.roars.it/online/?p=8216>>. (Cit. a p. 175).
- Reay, Hannah e Jackie Sears. «A collaborative model for training clinical research staff». *Nursing Management* 20.3. (2013): 22–27. (Cit. a p. 177).

- Regha, V Sakthi e BS Swaroop Rani. «Job psychograph of library and information science professionals in higher education institutions in Tamil Nadu». *Annals of Library and Information Studies* 53.1. (2006): 7–14. (Cit. a p. 179).
- Schechter, Neil L., et al. «Pain reduction during pediatric immunizations: evidence-based review and recommendations». *Pediatrics* 119.5. DOI: [10.1542/peds.2006-1107](https://doi.org/10.1542/peds.2006-1107). (2007): 1184–1198. (Cit. a p. 202).
- Sen, BK. «Ranganathan's five laws». *Annals of Library and Information Studies* 55.2. (2008): 87–90. (Cit. a p. 179).
- Silvestro, Annalisa. «Infermieri: la forza di una nuova cultura per il sistema salute». *L'Infermiere: Notiziario Aggiornamenti Professionali, Organo ufficiale della Federazione Nazionale dei Collegi Ispasvi* 56.2. (2012): 8–20. (Cit. a p. 174).
- Singh, Surya Nath e B.S. Garg. «Impact of information technology (COMPUTERS) on biomedical information centres and libraries (ICLs) in India: a critical evaluation». *Annals of Library and Information Studies* 49.2. (2002): 51–66. (Cit. alle pp. 173, 176).
- Talab, Seyed Mohammad Ghaemi e Masoumeh Tajafari. «Impact of information and communication technology (ICT) on library staff training: A comparative study». *Annals of Library and Information Studies* 59.1. (2012): 7–15. (Cit. a p. 178).
- Thakur, DS e KS Thakur. «Approaches to knowledge management in higher education». *Annals of Library and Information Studies* 52.4. DOI: [10.1590/S0104-11692009000300017](https://doi.org/10.1590/S0104-11692009000300017). (2005): 115–118. (Cit. a p. 179).

MICHELE CHIEPPI, Fondazione I.R.C.C.S. Policlinico San Matteo di Pavia.

[biblioteca.laureaps@smatteo.pv.it](mailto:biblioteca.laureaps@smatteo.pv.it)

PIERA BERGOMI, Fondazione I.R.C.C.S. Policlinico San Matteo di Pavia.

[p.bergomi@smatteo.pv.it](mailto:p.bergomi@smatteo.pv.it)

DONATELLA DE AMICI, Fondazione I.R.C.C.S. Policlinico San Matteo di Pavia.

[ddeamici@smatteo.pv.it](mailto:ddeamici@smatteo.pv.it)

---

Chieppi, M., P. Bergomi. D., D. De Amici. "La ricerca bibliografica in infermieristica:

un'esperienza di condivisione tra strutture della Fondazione IRCCS Policlinico S. Matteo di Pavia". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art. #8922. DOI: [10.4403/jlis.it-8922](https://doi.org/10.4403/jlis.it-8922). Web.

ABSTRACT: With the aim of sharing strategies and bibliographic research skills, the IRCCS Policlinico S. Matteo di Pavia Foundation developed a training course titled: *Bibliographic research in nursing science*. Many offices contributed to the successful conclusion of this course: the Nursing library; the Training and Development office; the degree course in Nursing and the Technical and Rehab Nursing Service (SITRA) at the IRCCS Policlinico S. Matteo di Pavia Foundation. The two-years project was developed in two training levels, and it was addressed to nursing tutors and SITRA members. The two training levels helped in sharing the knowledge on functionalities and services offered by medical and nursing databases, and to apply complex research strategies. Participants' contribution helped in producing a bibliographic citation manual for nursing science.

KEYWORDS: Bibliographic research; Medicine; Nursing; Training.

---

Submission: 2013-06-13  
Accettazione: 2013-08-26  
Pubblicazione: 2014-01-01





# Research data and humanities: a European context

Maura Funari

## 1 Defining research data

The subject of research data<sup>1</sup> has handled in scientific research's contexts, both in theorizations on research's different topics, and in management policies of research's results and in international organizations policy's documents. To bypass the necessity of defining research data, in the strict sense of the word, is a common tendency. But virtually to meet with conceptual ambiguities is frequent: depending on the field of application, peculiar aspects are identified so they, even though in some cases almost imperceptibly, vary. So far one definition with a univocal and a universal kind, does not exist.

UNESCO (Swan) describes research data as a type of "research output" together with journals, peer-reviewed conference proceed-

---

<sup>1</sup>The italian translation has not found a clear definition: the italian version of European Commission's documents refers both to "dati della ricerca" (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell'investimento pubblico nella ricerca*) and to "dati di ricerca" (*Raccomandazione della Commissione, del 17 luglio 2012, sull'accesso all'informazione scientifica e sulla sua conservazione*). In this treatise the Anglo-Saxon term has been kept because of the common acceptance.



ings, and books. This categorization seems to individualize in this typology of data, an instrument to externalize research's products, also acknowledging to research data a general increase in the attention that open access policies attach them.

A set of definitions comes from Australian National Data Service (ANDS) which, specifying in the introductory the fact that any definition is likely to depend on the context in which the question is asked, collects definitions enunciated in data management policies of some Australian universities: University of Melbourne, Monash University, Griffith University (Australian National Data Service). The first two definitions quote common elements in research data's characterization, identifying a variety of shapes and contents and so leaving out a previous determination based on these criteria. The founding role of research data, that is the datum used as primary source, or on which the research's theory is based, is an additional aspect just in the first of the three quoted definitions: "Research Data means data [...] on which an argument, theory, test or hypothesis, or another research output is based" (University Of Melbourne). The Griffith University's definition verbatim refers, in the first part, to the OECD Organisation for Economic Cooperation and Development (OECD) one (13) in which research data are defined as "factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research and that are commonly accepted, in the scientific community, as necessary to validate research findings". Therefore this definition determines that the condition of commonly acknowledged, as necessary material in order to validate research findings by scientific community, is determining for research data. A similar formulation comes from definition of scientific data reported in the Memorandum for the Heads of Executive Departments and Agencies<sup>2</sup> of the Executive Office of the Presi-

---

<sup>2</sup>[http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_)

dent, Office of Science and Technology policy, of the White House. These data are defined as: “the digital recorded factual material, commonly accepted in the scientific community, as necessary to validate research findings including data sets used to support scholarly publications”. Also in this case the condition of necessary material to validate research findings, not defined as “factual record” but as “factual material (digital recorded)”, is connotative.

The ANDS is, moreover, the RDA’s founder<sup>3</sup> (Research Data Alliance), with the European Commission’s support (through the iCordi project, treated in the paragraph assigned to infrastructures) and the United States one too (through the National Science Foundation). The international organization aims to accelerate and to improve innovation and data-driven research, encouraging the actions connected to research data (such as exchange, sharing, uses and re-uses, standards and visibility) and achieving the development and the adoption of infrastructures, policies, practices, standards e services.

Getting back to the research data’s definitions, the Communication of the European Commission (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell’investimento pubblico nella ricerca* 3), emphasizing the increasing attention to the research data’s access improvement, characterizes them as “experimental results, observations and computer-generated information which form the basis for the quantitative analysis underpinning many scientific publications”. Holding into consideration the heterogeneity coming out of the above-said definitions, however, it is possible to establish that research data can be meant to data, in different forms and contents,

---

public\_access\_memo\_2013.pdf.

<sup>3</sup><http://rd-alliance.org/>.

which constitute the basis for a scientific research, as a primary resource and foundation of the research's findings. Since the intrinsic value and the necessity of collection, preservation, sharing, are qualities varying according to different factors, from the nature of the research to the disciplinary field, the individuation of pre-established criteria is essential to set what kind of data, produced into a specific area, could be included in research data's category. This task should mainly belong to research's findings treatment policies developed by centers, agencies and institutions involved.

From both a qualitative and a quantitative point of view, during the research's phases, a lot of data could be produced but certainly their potential value constitutes the essential assumption for the interest toward this typology, in the fields of information science and research's findings treatment policies. This potential value can vary depending on datum's form, nature, origin (National Science Foundation. National Science Board12-13). This relativity is sharpened by differences that emerge both in natural sciences/human sciences macro-areas, and in the single disciplines, inside the two areas.

Furthermore it must be specify that research's datum can has got the double role of product (as result or resultant of a specific research) and of source (as a datum already produced by someone else and re-used as the basis of a new research): a circumstance that has been appointed referring to the antithetical ideas of output and input.<sup>4</sup> This doubleness brings out the pattern of a circulation and knowledge's sharing system which, whereas an open level of sharing is looming, founded on the action of re-using (Murray-Rust)(Murray-Rust). Murray himself, quoted in the Italian studies (De Robbio and Giacomazzi), notices one difference in practices of data's publication

---

<sup>4</sup>"Data are outputs of research, inputs to scholarly publications, and inputs to subsequent research and learning" (Borgman, *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*115).

and use between the typologies of “Large Science” and “Small Science”.<sup>5</sup> Furthermore, from a terminological point of view,<sup>6</sup> the open level of sharing is the discriminating condition between research data and open research data: the last one refers just to open data, while the first does not leave out them.

## 2 Research data in humanities

Natural sciences differ from humanities, besides for field of study and methodologies, also for a greater quantity of data produced (as a consequence of the technical level and the objectivity founding natural sciences), for their typology (which affects also their level of elaboration) and also for the degree of necessity and practice in their sharing and re-using. These circumstances have made scientific-natural research’s data protagonists, not just in theoretical studies, but also in the executions of systems for their collection, management and sharing, and in the policies concerning to research. The disadvantage characterizing humanities in this field, such as in the open access one (Suber), derives both from scientific-humanistic research’s features and so from their results and sources, and from correlated economic and cultural explanations, especially concerning spread and timing (in the scientific-natural research the need of sharing, both in results ’s storage and in the access to them, is characterized by a quickness that is reduced in humanities.)<sup>7</sup>

---

<sup>5</sup>The distinction made by Murray between “Large science” and “Small science” is based on the research’s unit dimension that, in the first case, is vast and narrow (individual or laboratory) in the second one.

<sup>6</sup>The requirements satisfying the attribute of Open, referring to data are, various. See the definition of “open” proposed by the Open Knowledge Foundation: <http://opendefinition.org/okd/>.

<sup>7</sup>Peter Suber (“Promoting open access in the humanities”) in his analysis on the open access slow moving in the humanities compared to natural sciences, identifies

If according to natural sciences the individualization of what can correspond to a research data is rather intuitive, for when it comes humanities, the question is more complex.<sup>8</sup> The National Science Foundation (12-13) identifies different data categories (“observational, computational, or experimental”) as a result of their giving origin to an observation, a computation, or an experiment. If in natural sciences these actions belong to the standard researching method during the interpretation and the study of phenomena, understood as observational events, in humanities, which have as subject-matter no physical entities, the correspondence is not so obvious.

The overlapping between data concept and “primary source” (Burrows) constitutes the key point of the matter.<sup>9</sup> The connection between the two concepts is clear in the OECD’s quoted definition (*Principles and Guidelines for Access to Research Data from Public Funding*), in which research data are the “factual records” used as “primary sources”. This role comes true in the research’s process: the datum considered as factual record becomes the primary source of the research. And yet, considering the data products and their

---

nine differences concerning the research in the two different areas. Although the analysis’s subject-matter are the research’s findings in the form of journal articles, many observed circumstances are valid also for when it comes research data, outlining a contest in which open access in humanities appears as “less urgent and harder to subsidize, than in the sciences”.

<sup>8</sup>In Borgman (“The digital future is now: A call to action for the humanities”) the question “What constitute data in the humanities?” is investigated but it does not found a clear answer. The author concludes, referring to the quoted question and to other four questions concerning digital humanities: “Answering these questions will enable the digital humanities community to be more articulate about its scope and its goals, and better positioned to identify their requirements for infrastructure”. (The question is broached also in Borgman, *Scholarship in the Digital Age: Information, Infrastructure, and the Internet* 215-217).

<sup>9</sup>Burrows (“Sharing humanities data for e-research: conceptual and technical issues”) claims that to not discern “primary source” from “data” in humanities “would be analogous to describing the stars and galaxies as an astronomer’s ‘data’”.

utilization in the scientific humanistic research as factual records is hard, even if this circumstance can not be excluded at all, referring for example to the findings of a text mining in textual analysis, or to archeographic data in the archaeological field. However, it is beyond doubt that, from a quantitative and a qualitative point of view, this data typology does not represent the pre-eminent one, also taking into consideration the nature of investigation's subject, often constituted by abstract entities in the form of representation. Instead of simple data, in most cases, they are data sources (data and information's sources), in different forms, varying from texts to objects.

Briefly said, to identify what to mean for research data in humanities, taking into consideration in a unavoidable way, the discretionary power conferred from the "primary" founding label which characterizes research data's definitions and role, it is necessary to make a distinction between:

- Data intended as immediately knowable elements, as a result of observations of phenomena, reality, experiments, computations. (Examples: findings of a text mining in textual analysis, archeographic data, questionnaires, audio gatherings in field researches, etc.).
- Data sources as investigated elements providing the datum, or on which the datum itself is based. (Examples: painting, literary work, musical manuscript, archaeological find, etc.).

This last typology, being on the basis of the datum, represents its source, where the research leads to findings belonging to the first category (which constitute the proper research data). For what concerns the production and the use in the scientific-humanistic research, the first typology, as said before, is not quantitative and qualitative substantial as in scientific-natural one, in which the study

of physical phenomena generates a computational or not production of raw data, directly recorded. Otherwise the using of resources, which are constituted by both data and sources (ex. documents, images, texts, etc.), is relevant so much so that, as explained in the dedicated paragraph, the European infrastructures expressly collect digital resources. In the humanities area can not be said that the factual records are equivalent to the primary sources and, therefore, that these constitute the research data, but it is possible to detect that they have got the same role, as basis of a research. Furthermore, in humanities, technology offers in many cases the direct possibility of recording, reproduction, graphic representation, accessing and linking<sup>10</sup> (circumstance less likely in the natural sciences area). So, digitalization, if integrated with accessible infrastructures, makes possible the collection, the sharing and the use of resources' collections, even if leaving out their materiality.

### 3 Research data in European policies

The attention to research data, among European Community, has been realized through a series of European Commission's Communications (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio e al Comitato economico e sociale europeo. Sull'informazione scientifica nell'era digitale: accesso, diffusione e conservazione comunicazione; Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Le Infrastrutture TIC per la e-scienza; Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. Un'agenda digitale europea; Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al*

---

<sup>10</sup>An example of digitalized data sources collection, linked to data, is Europeana [www.europeana.eu](http://www.europeana.eu).

*Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell'investimento pubblico nella ricerca*), till it has reached the form of a Recommendation (*Raccomandazione della Commissione, del 17 luglio 2012, sull'accesso all'informazione scientifica e sulla sua conservazione*) to member States. This interest is aim to research data and, more in general, to scientific information, produced in all research's fields, among which humanities are quoted.<sup>11</sup> Research data, as a parallel form but also as associated<sup>12</sup> to the proper publications, constitute the scientific information, of which wide and quick diffusion has a central role in terms of research's innovation, progress, efficiency and excellence. The desirability of this diffusion, however valid on principle, is necessary especially in the case of findings, both journal articles and data, resulting from publicly funded researches.<sup>13</sup>

Already in 2009, in the field of ICT Infrastructures (Information and Communication Technologies) for e-science, the assumption of the emergence of new research methods that exploit advanced computational resources and data collections, as well as the awareness

---

<sup>11</sup>"The emergence of 'big data science' has a global dimension, as it reflects the increasing value of raw observational and experimental data in virtually all fields of science (humanities, biodiversity, high-energy physics, astronomy, etc.)" (Commissione Europea, *Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Le Infrastrutture TIC per la e-scienza*9).

<sup>12</sup>The Communication deals with a "'continuum' of the scientific information space from raw data to publications across different communities and countries". Internet and the new information and communication instruments allow, indeed, to use research data coming from experiments and observations, associating them to other information's sources, to the aim of taking out meanings (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio e al Comitato economico e sociale europeo. Sull'informazione scientifica nell'era digitale: accesso, diffusione e conservazione comunicazione*3).

<sup>13</sup>*Ibidem* However some delay for the first use by researchers or for commercial purposes can be considered as justifiable (3).

of “the increasing value of raw observational and experimental data in virtually all fields of science” (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Le Infrastrutture TIC per la e-scienza*<sup>9</sup>), had identified as objective, the adoption by Europe of a “coherent and managed eco-system of repositories of scientific information” (11). Member States and scientific communities had been asked to step up investments in scientific data infrastructures, as also provided for by the Commission into the Seventh Framework Programme, with the aim of “support accessibility and preservation policies”. In this ambit of interest toward research data takes place the Communication (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell’investimento pubblico nella ricerca*) followed by the Recommendation (*Raccomandazione della Commissione, del 17 luglio 2012, sull’accesso all’informazione scientifica e sulla sua conservazione*), in which the attention to them becomes more detailed. The traditional debate focused just on publications and, at the same time, the increasing importance of the improvement of access to research data, defined as already quoted, are indeed acknowledged. The inefficiency of public research investments reveals itself where findings in the form of data, for the verification and the possible use, are made not available to a wide public of users. In the Communication are identified different obstacles to the development of this new clutch of knowledge sharing and are stated the initiatives already carried out by the Commission (OpenAIRE) and those planned (financial supports to data infrastructures and to research on digital preservation). The identified obstacles related to the development of research data access and to their use and re-use are:

- The lack of organization and clarity about responsibilities.

- The lack of financing models to ensure long-term access.
- Persistence of interoperability problem among countries and disciplines.
- Researchers and innovative enterprises reluctance due to different reasons (data perceived as their property, time to invest in the practicalities of depositing, absence of reward and recognition mechanisms, such as citation mechanisms and measurement of the data citation impact (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell'investimento pubblico nella ricerca*<sup>7</sup>)).

Although these problems are real, other and more specific (here omitted) factors affect negatively on the process. It's about barriers<sup>14</sup> closely related to the access and to the use of research data, in the field of dedicated infrastructures. These barriers are legal (copyright, restrictive licenses, limitative editorial policies), financial (subscription to datasets 's access, payment for the use of materials) and technical (restricted visibility, restricted length, impossibility of accessing/using/re-using). Their existence and the weight held in the difficulty of realization and in the efficiency of systems for free and open research's findings sharing in the form of data, are due to the commercial value which often many data have, to the acknowledgment of creative work qualification, and so susceptible to the copyright, to the absence of sustainability plannings aimed to maximize the investment and to guarantee long term effects.

The importance assigned to research data comes out also into the new Framework Programme for Research and Innovation (2012-

---

<sup>14</sup>On the definition of the different kind of barriers see Murray-Rust ("Open Data in Science") and Suber ("Promoting open access in the humanities").

2020) called Horizon 2020.<sup>15</sup> The Proposal for decision (*Proposta di decisione del Consiglio che stabilisce il programma specifico recante attuazione del programma quadro di ricerca e innovazione (2014-2020) - Orizzonte 2020*) of European Commission Council dealing with the realization of the quoted programme, identifies in research's findings dissemination and communication on a continent wide, a "key added value" in order to enhance their impact (22). In line with this awareness are collocated the actions aimed to support the creation, the development and the operation of the TIC infrastructures with the goal to "achieve, by 2020, a single and open European space for online research where researchers enjoy leading-edge, ubiquitous and reliable services for networking and computing, and seamless and open access to e-Science environments and global data resources" (36). Research in social sciences and humanities is declaredly fully integrated in this specific objective concerned research infrastructures and in those generals of the programme (21). So, in Horizon 2020, are provided actions aimed to realize open access to research data. Indeed, the Communication provides for the launching of a "pilot scheme on open access to and re-use of research data generated by projects in selected areas of Horizon 2020." (*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell'investimento pubblico nella ricerca*9). The Recommendation exposes the urgency of adopting political actions on access to data and, therefore, recommends to member States to define clear policies providing for objectives and indicators to measure progress, implementation plans and financial plannings, to guarantee that "research data that result from publicly funded research become publicly accessible, usable and reusable through digital e-infrastructures" (*Raccomandazione della*

---

<sup>15</sup>[http://ec.europa.eu/research/horizon2020/index\\_en.cfm..](http://ec.europa.eu/research/horizon2020/index_en.cfm..)

*Commissione, del 17 luglio 2012, sull'accesso all'informazione scientifica e sulla sua conservazione*<sup>3</sup>).

The European Commission approach in defining research data policies is characterized, as already seen, by an attention to scientific information on the whole. Research data, in case open and accessible, constitute one of the instruments (with the other results) to strengthen the research system, in an interoperable and cooperative European context, both inside and in the extra European ambit. This setting out, which practically leaves out investigations and analysis on research data as an independent element, derives both from the nature of Commission's documents (they are not guides or technical texts) and from the propulsive role that the Commission has got, as a coordination center between the members States, for when it comes the development and the improvement of the scientific research system on the whole.

A different setting out is presented by the OECD's contribution (*Principles and Guidelines for Access to Research Data from Public Funding*), in a guide form, and addressed to the single States to encourage an international and efficient research data sharing and use, overcoming variety of laws, policies and national practices. The text offers an analysis of the different aspects concerning the definition of policies by research's institutions and founding agencies. Furthermore, as seen, it offers a clear and complete definition of this data category, placing it in a separated dimension, however not subordinate or necessarily parallel to the other research's results. After an introduction specifying that the principle of opening and ideas, information and knowledge free sharing is on the basis of Organization member States scientific public systems, is recognized that new technologies have created a "new fields of application for not only the results of research, but the sources of research: the base material of research data" (9) and that an "effective" access to these data, giving enun-

ated benefits and advantages, should improve profits coming from public investments. The subject of this access are data resulting from public founding researches. The known advantages resulting from these data open sharing have a character both universal and individual for the single members States. The principles on which data research have to be founded (openness, flexibility, transparency, legal conformity, protection of intellectual property, formal responsibility, professionalism, interoperability, quality, security, efficiency, accountability, sustainability) are, as said, enunciated and analyzed in their different aspects.

Also the UNESCO (Swan) stated the research data value, defining them, as seen, as a "research output" category which is increasingly getting a central role in the open access policies. The text, which has the aim to promote the open access in the member States, facilitating the understanding of the connected aspects, is concentrated on the relevant issues. Even though the research data belonging to research output group (so to the research information on the whole) and the inclusion of this research's results typology in the open access concept are recognized, open access main and original "target" is determined in the "journal literature" (10). It's just this category the main subject of the work. In the document, however, is underlined the centrality of data-intensive sciences in the open sharing process of finding research and the differences of rules and data management between disciplines. For when it comes strategies to promote open access is admitted an increasing difficulty in separate open access to the "literature" from open data (referred to research data), and the consequent need to include in future strategies supporting open access, those concerning data. Although this accepted tie<sup>16</sup> the need to develop diversified policies for open data, that considering problems connected to privacy and circumstances

---

<sup>16</sup>Swan (27) talks about an "ecosystem of 'open' issues".

preventing the diffusion because of other reasons, is highlighted.

A policy on the research results access is modelled on the subject it regulates: in defining contents, storage methodologies, standards and everything else, subject's characteristics of the collection, preservation, sharing and re-using process can not be neglected. The product, in the datum form, takes a different shape from a journal article. Some aspects related to the quoted category need, in the policies 's definition, a particular attention: for example, data can be contained confidential information (such as data collected in social sciences field) or can have a commercial value. Therefore, in these cases it is necessary to identify provided actions aim to establish conditions and specific limits. The characteristics of research data differ not only from those of scientific articles (for this reason a separated policy for the two research output categories would avoid omissions and gaps), but also from the single research ambits. Indeed it has been gathered that "a generic approach to data curation will not be sufficient to cope with the different data-related needs and expectations of researchers working in different disciplines other than at a superficial level" (Key Perspectives2).

Although the variety of contexts in which research data theme is broached, in European policies, involving a difference of purposes and approaches, the research data value (related both to scientific-humanistic research and to scientific-natural one), guided towards the improvement of the research's process and of everything else related to, can be said unanimously recognized. This value materializes in the open access to data through TIC infrastructures, with the aim to make them widely accessible, usable and re-usable.

The presuppositions to the necessity of making data widely available are valid on principle, since results from public funding projects and researches should return to the funding community. In this way, not only citizens 's right to potentially take advantage of final findings is

observed, but also the investment itself is optimized: the researches 's repetition is limited (if data derived from previous researches are not available, it is necessary to repeat the research), times for research are increased thanks to the speed in information's finding (also at an interdisciplinary level), costs for the access are removed (for example, the access to some datasets accessible for payment), the research system on the whole becomes more transparent (data quality can be verified and the datasets 's use is measurable). All that positively influences, in a more or less direct way, the growth, the progress and the development at an economic and social level (For a detailed list of the advantages in terms of public investment optimization in the scientific research, see Organisation for Economic Cooperation and Development<sup>10</sup>).

## 4 The Italian position

To make research data value concrete in a practical form, need systems aimed to collection, management, preservation, sharing, exploitation of research's findings (and everything connected to these operations) and States 's action to support and to promote these initiatives. Since they are real "systems", by definition corresponding to sets of instruments, mechanisms and elements, the coordinative action is essential, both to an internal level and to an external one. As it can be intuitively understood, also the financial aspect is included in the actions for these systems, together with the spread of a cultural education about research's datum. Currently, Italy faces two circumstances which constitute the starting point of the implementation of policies, focused on the quoted actions: the first one is a political constraint deriving from European Commission Recommendations and Communications, for which, as a part of an aggregative organism, should conform to common policies; the

second circumstance is a practical aspect constituted by advantages in terms of cultural, economical and social progress, coming from development, improvement, growth, exploitation, efficiency of the scientific research process. In this condition are fitted in the participations to European intra and extra initiatives in favour of the development of research data collection, management, sharing systems. (ex. CLARIN, DARIAH, ARIADNE, MAPPA, RDA and iCordi, seen in the next paragraph). Compared to other European countries, which are in the lead for when it comes initiatives to support research data projects, such as the United Kingdom one, it is clear that Italy is in a backwardness position. First of all, national policies aimed to discipline the matter, in terms of definition of the different aspects connected to the spread of research's results (from the contents identification to the responsibilities) lack; there is not a national reference point as a coordination centre; financial and organizational models, assuring the infrastructures 's long term sustainability, have not been outlined.<sup>17</sup> With regard to existent infrastructures, besides the limit constituted by the lack of a multiannual planning aimed to guarantee the long term sustainability, there are a limited capability to exploit the social-economical benefits connected to the realization and to the operation of a research's infrastructure and an unfit presence of European infrastructures into the national territory. The quoted participation to European infrastructures projects is real, but properly defined as "assicurata da gruppi di ricerca di massa subcritica" (Ministero dell'Istruzione, dell'Università e della Ricerca<sup>66</sup>): so it is not sufficient to set off a widespread phenomenon. The Horizon

---

<sup>17</sup>Conservations' structures are often created to specific projects and the fundings are limited to a certain period (For when it comes the obstacles to research data access and to their use and re-use and to the long term preservation, see Commissione Europea, *Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell'investimento pubblico nella ricerca*<sup>7</sup>).

2020 Italia document, drawn up by the MIUR (8), on communication of the research and its results deals with italian “limitata capacità” of “trasferimento, diffusione e valorizzazione” depending on both italian researchers, and still insufficient support services of universities and national Public Research Agencies. Furthermore, the same document underlines the existence of “criticità importanti per il posizionamento del sistema europeo della ricerca e dell’innovazione e, al suo interno, in maniera più accentuata, di quello italiano” (5). In accordance with the comunitary programme, Italy intends to realize “un sistema della ricerca sano ed efficiente, non frammentato e privo di duplicazioni, forte, coeso e strategicamente orientato” (26). To pursue this scope, the circulation and the sharing of scientific research’s results have a relevant position. Indeed, the document states the importance that the access (open, free and with an interoperable format) to data and information resulting from public funded activities has, with regard to the connection between science and society and to the optimization of the financial investment itself. The Researchitaly portal<sup>18</sup> is identified as the aggregating gateway of the initiatives on open access of the national research system and as the platform to list local repositories of universities and research’s centers and to contain a national repository. For the Italian system of research’s infrastructures in which, as quoted, are highlighted a series of lacks, is provided a national plan (PNIR), aimed to improving them. Moreover, is planned the strengthening of the existent infrastructures and the realization of new ones, according to the European Strategy Forum on Research Infrastructures (ESFRI) and the legal instrument European Research Infrastructure Consortium (ERIC)<sup>19</sup> is identified as valid to take part in pan European infrastructures projects.

---

<sup>18</sup><https://www.researchitaly.it/>.

<sup>19</sup>[http://ec.europa.eu/research/infrastructures/index\\_en.cfm?pg=eric1](http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric1).

The perspectives enunciated in Horizon 2020 Italia programme, reveals a new scenario for the country: if it becomes reality, we will be a European country that eventually will takes part actively in knowledge's free movement.<sup>20</sup>

## 5 European infrastructures

The increasing attention to research data, also in humanities, besides theories, has been carried out in European projects and infrastructures, both with an international participation and with individual country or centres initiatives, collecting and making available to access these data, or providing the instruments to support and to promote these initiatives. In many cases the process is integrated in a context of interest toward digital resources concerning the research, in a general and widened meaning. In the practice, indeed, the distinction between research data and data is not clean: the firsts are a subcategory of the seconds, and being the infrastructures, as said, primarily aimed to collect digital data or resources resulting from research, often also data sources are included in the data collections, together with research's results intended as raw. This circumstance causes a difficult individuation of resources 's nature really preserved and accessible in the different infrastructures and the almost impossibility of a clear and certain classification.

Repositories and research data list published by DataCite,<sup>21</sup> re-

---

<sup>20</sup>Referring to the knowledge's free movement, Janez Potocnik, European commissioner for the Science and the Research, in 2007, in the occasion of the Green Book presentation «The European Research Area: New Perspectives», talked about a "Fifth freedom". The other four freedoms are those of common market (free movement of people, services, products and capital) enunciated in the CEE Treaty (1957). For more information, see: [http://cordis.europa.eu/fetch?CALLER=NEWSLINK\\_IT\\_C&RCN=27454&ACTION=D](http://cordis.europa.eu/fetch?CALLER=NEWSLINK_IT_C&RCN=27454&ACTION=D).

<sup>21</sup><http://datacite.org/repolist>, the list is continuously updated.

alized through the DataBib initiative<sup>22</sup> (the instrument for the individualization and location of these infrastructures), shows that, between European countries, most of humanities repositories is located in the United Kingdom. For when it comes other countries such as France, Sweden and the Netherlands, the result is just one for each country.

The Registry of research data repositories re3data.org<sup>23</sup> founded by the DFG (German Research Foundation) allows to search by repositories 's subject and to the *Humanities and social sciences* category are linked nine results. Most of these (with the exception of two cases) are European repositories. Between the individual States, a well-advanced position in the realization of digital resources infrastructure, including research data, is owed, as said, to the United Kingdom.

The Arts and Humanities Data Service (AHDS) service<sup>24</sup>, founded by JISC Joint Information Systems Committee (JISC) and JISC, was born in 1996 as a national service, with the aim to collect, preserve and promote electronic resources resulting from research and teaching in the arts and humanities. Funded until the end of March 2008, is now decentralized between host institutions. Divided into five disciplinary areas (archaeology, history, literature languages and linguistics, performing arts, visual arts), it incorporates the infrastructures sharing the same scopes, collecting and making accessible, those defined digital resources trough data archives. The Archaeology disciplinary area is hosted by Archaeology Data Service (ADS),<sup>25</sup> founded by a consortium constituted by the Council for British Archaeology and the Universities of Birmingham, Bradford, Glasgow, Kent at Canterbury, Leicester, Newcastle, Oxford and York. The

---

<sup>22</sup><http://databib.org>.

<sup>23</sup><http://www.re3data.org>.

<sup>24</sup><http://www.ahds.ac.uk>.

<sup>25</sup><http://archaeologydataservice.ac.uk>.

aim of the service is to collect, describe, catalogue, preserve and provide user support for “digital resources that are created as a product of archaeological research”. Moreover, the ADS promotes standards and guidelines for best practice of creation, description, preservation and use of archaeological information. In collaboration with national and local agencies, involved in the funding of archaeological research or in the funding of archaeological research (Arts and Humanities Research Council, Natural Environment Research Council, British Academy, Council for British Archaeology, English Heritage, Society of Antiquaries of London), it collects datasets of different origin and typology and resources from maps to text report. The research’s mechanism, for the user, is structured in two different systems, both much framed. Archsearch for the records allow interrogations by key words, preset categories (What, Where, When) and resources (intended as collections). In the system called Archive the research is set by archives classified into subject, programme and region. Furthermore, there are additional features which not only allow definite research’s narrowing, but also offer advanced instruments, such as the map function and the external research (at the moment in an experimental phase).

A French national platform, launched in December 2010, funded by the government, which collects, enriches and provides a unified structure of access to digital data in humanities and social sciences is Isidore.<sup>26</sup> Created by TGE ADONIS (nowadays merged in the TGIR Huma-Num) and realized by the Centre pour la Communication Scientifique Directe (CCSD) with the participation of the Antidot, Sword e Mondeca societies, it is defined research platform and currently collects 80 collections, 2026 sources and 2.271.736 resources.<sup>27</sup> The quick interface allows a prompt research by key word,

---

<sup>26</sup><http://www.rechercheisidore.fr/index>.

<sup>27</sup>Figures recorded in May 2013.

by categories (type of resource, category, historical period, discipline, collection etc.), by sources and by repositories.

The TGIR Huma-Num, in a section dedicated to the project in its web site,<sup>28</sup> tackles the question of the peculiarity of humanities and social sciences in relation with the need of data produced sharing and collection. Acknowledging a central role to the information's sources ("sources d'informations"), in particular, is highlighted that, between these, the text is basic in the knowledge production process in humanities and social sciences. The time, that often makes a scientific text rapidly obsolete, does not make the same for when it comes a medieval manuscript, which can preserve the same importance and topicality as the last article published in an international journal.

Some international collaboration infrastructures offer support and services with the aim of improve research's results access and use in humanities, realizing real networks: an example is Digital Research Infrastructure for the Arts and Humanities (DARIAH)<sup>29</sup> that currently counts fourteen member States, including Italy. The project's origins date back to 2005 but the preparatory phase, aimed to identify and define the elements (physical, strategical and human) and financial and legal aspects, concluded in February 2011. The declared aim of DARIAH is enhance and support digital research in humanities and in the arts, realizing a cooperative infrastructure, that puts together national, regional and local efforts, in an interconnected network of instruments, people, information, methodologies etc. DARIAH's has been created in the ERIC form and this normative frame facilitates the long term sustainability of the project: financial and technological needs are confronted in collaboration between members, in a coordinated and uniform atmosphere. The project's grand vision is aimed to facilitate long-term access and use

---

<sup>28</sup><http://www.huma-num.fr>.

<sup>29</sup><http://www.dariah.eu>.

for all the “European Arts and Humanities (A+H) digital research data”. DARIAH operates through four virtual competency centres, the Virtual Competency Centres (VCCs), each of them operating in a specific area (e-Infrastructure, Research and Education, Scholarly Content Management, Advocacy).

Common Language Resources and Technology Infrastructure (CLARIN),<sup>30</sup> exclusively dedicated to linguistic resources, is another pan European infrastructure with members and institutions from thirty-three countries, which offers services to provide a easy access to the resources, through an integrated and interoperable system. Specifically, the scope of the project is to encourage the research’s progress in humanities and in the social sciences realizing a unified single sign-on access platform which integrates, at a European level, language-based resources and advanced tools, creating a shared and distributed infrastructure. The preparatory phase of the project started in 2008 and ended in 2011; nowadays CLARIN is still under construction, but a set of services are available and accessible. These are divided into two typologies and the first is constituted by the services for users relating to research, transformation, resources’s archiving: the depositing service, the Virtual Language Observatory (VLO), web services and consulting services. The second typology concerns technical infrastructure services, for CLARIN centers: CLARIN IdP (Identity Provider), CLARIN Discovery Service, Component registry for the Metadata Infrastructure (CMDI), ISOcat concept registry, Relation Registry.

With regard to infrastructures dedicated to scientific data, the iCordi project<sup>31</sup> (which through the European Commission, as said, supports the Research Data Alliance) represents an important goal, most of all in terms of practical attention to interoperability. Started

---

<sup>30</sup><http://www.clarin.eu>.

<sup>31</sup><https://www.icordi.eu>.

in September 2012 and funded by European Commission, into the Seventh Framework Programme, the project pursues the aim to realize a coordination platform between Europe and the USA, to discuss and improve the interoperability of scientific data infrastructures and to extend this to the global level. In the specific instance, the declared strategic vision is to give an important contribute to the development of the policy for the management and the curation of scientific data, leading to a common policy addressed to the development of a wider global infrastructure. ICordi, which currently counts fourteen partners, will be guided by an High Level Scientific Forum composed by experts in the fields of management and curation policy, coming from both the involved continents, which will realize strategic recommendations aimed to improve the concurrence of data integration, interoperability and infrastructures. The action will be based on three programmes (analysis, prototype and workshop): the first will be devote to analyze data organizations and solutions, as they emerge from the various scientific communities; the prototype one will coordinate activities between Europe and United States important projects, supporting cross-infrastructure experiments on EU-USA interoperability; the workshop programme will investigate the infrastructures 's convergence, paying a particular attention to a wide set of scientific disciplines.

For when it comes individual disciplines, important developments (and almost just italian initiatives)<sup>32</sup> has been realized in archaeological field, as a consequence of the subject's characteristics which, because of needs and methodologies, make research data sharing an exigency, besides a value added. The Advanced

---

<sup>32</sup>Italy, through institutions and centers, takes part in European infrastructures projects, such as CLARIN and DARIAH. In the archaeological field the Metodologie Applicate alla Predittività del Potenziale Archeologico (MAPPA) has realized MOD (MAPPA Open Data, an archaeological digital repository, available at <http://mappaproject.arch.unipi.it/mod/Index.php>.

Research Infrastructure for Archaeological Dataset Networking in Europe (ARIADNE) project<sup>33</sup> launched in February 2013, with the expected duration of four years and funded by the European Commission, into the Seventh Framework Programme, aims to the creation of an archaeological data infrastructure. The scope is to realize an infrastructure for a transnational access to data, to instruments and to guidelines through a collection of several databases, offering a unified access point and instruments to place new technologies at research's disposal. Coordinated by the PIN (Polo Universitario Città di Prato) of Università degli studi di Firenze, with the collaboration of the Ministero per i Beni e le Attività Culturali and other Italian institutes, it gathers partners coming from sixteen European countries.

This brief roundup of some main examples of research data infrastructures in humanities, shows that Europe proceeds in two different but analogous and linked ways. Indeed, on the one hand, there is the creation of international level infrastructures that collect, support, address, answering to a need of coordination and collaboration; on the other hand, the subject of this coordinative action seems to be the process of creation and need of research data sharing, collected in other infrastructures through independent or not initiatives. The single States role in defining policies to promote, support, improve, enhance this process in progress, is crucial not just to make the individual enterprises efficient, but also to restrict lacks of homogeneity, favoring the realization of systems easily identifiable and usable singly and integrated.

---

<sup>33</sup><http://www.ariadne-infrastructure.eu>.

## 6 Conclusions

Research data constitute a typology of material which have a value that, in the field of research and in the more general sharing of knowledge, enables it to have a place in the European policies and in the international organizations' documents. The production and the existence of this kind of data, integrated with the possibilities offered by information technologies (systems of collection, management, preservation, interoperable languages and formats, the internet), have led to the awareness of the advantages, at different levels, coming from their free sharing (and from their free use and possible re-use). In particular, for research data resulting from public funding researches, to the reason of advantages given to science, progress, economy, society, is added the one concerning the necessity (if not the duty) of making research's results available to the same funding users. If in theory (how it can be inferred also by documents of the European Commission, the OECD and the UNESCO) all that is recognized and, in the case of the European Commission, the awareness is addressed to real initiatives, in practice difficulties exist and persist. The obstacles to the realization, to the development and to the efficiency of systems for free and open sharing of research data are of different kinds: financial, organizational, technical, legal, cultural. To identify them and to find out answers and solutions is crucial to realize valid and solid models and infrastructures aimed to guarantee the correct collection, diffusion, preservation and sharing of research data. However, until then, data will continue to be produced but many of them will not be available, other will be available for restricted lengths of time, other more will not be available for everybody (because of circumstances depending on, for example, visibility, formats or access conditions). This will make investments inefficient (or efficient in a reduced manner): many researches will be repeated, times for the finding of material might be

long, resources invested in projects without adequate financial and risk planning could be lost. However, the European situation, on the whole, seems to move toward the realization of policies, systems and infrastructures dedicated to research data, paying attention to what guarantees the efficiency (financial and technical sustainability, interoperability, etc.); see the examples of DARIAH and CLARIN projects.

In Italy the purpose of tackling the research data question emerges, for the first time, in the document Horizon 2020 Italia. So that this declared intention would not remain just a purpose but would become a concrete reality, some State's actions will be essential:

- In the scope of responsibilities 's individualization, the realization of a national coordination center in the field of open access to resources resulting from italian research.
- Financial investments into national initiatives and participation to pan European ones, such as for example the realization of a national infrastructure, integrated with European systems.
- Promotion of an open access culture, that would remove all fear (often springed from the lack of knowledge of new diffusion models and practices, such as the Creative Commons licenses), proposing new solutions.
- Complete and clear policies 's definition, that maintain the separation between open access to scientific articles, and to research data. Moreover it is important to maintain a conceptual separation between the open data in a generic sense (which include, for example, data from public administration) and research data which have proper and specific characteristics and questions and which need focused actions. These policies should draw up multi-year financial plannings, but also strate-

gic ones (such as risk action plans), with the aim to guarantee the long term sustainability.

Finally, it would be suitable and desirable that the question of research data in humanities, which, as said, suffer from the subordination as regards attention they have in scientific-natural sciences, be considered with the peculiarities which characterized the research and its results in this ambit. To look at the reality of humanities paying attention to the peculiarity that characterizes it, would stimulate a greater interest and development, keeping the value of the comparison and the sharing of solutions, practices, questions and attributes between the two sectors.

## References

- Australian National Data Service. *What is research data?* ANDS, 2011. <http://ands.org.au/guides/what-is-research-data.pdf>. (Cit. on p. 210). Print.
- Borgman, Christine L. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge: MIT, 2007. (Cit. on pp. 212, 214). Print.
- . “The digital future is now: A call to action for the humanities”. *Digital humanities quarterly* 3.4. (2009). (Cit. on p. 214). Web.
- Burrows, Toby. “Sharing humanities data for e-research: conceptual and technical issues”. (2011). (Cit. on p. 214). Web.
- Commissione Europea. *Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Le Infrastrutture TIC per la e-scienza*. Bruxelles, 2009. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0108:FIN:IT:PDF>. (Cit. on pp. 216–218). Print.
- . *Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. Un’agenda digitale europea*. Bruxelles, 2010. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:IT:HTML>. (Cit. on p. 216). Print.
- . *Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni. Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell’investimento pubblico nella ricerca*. Bruxelles, 2012. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0401:FIN:IT:PDF>. (Cit. on pp. 209, 211, 216, 218–220, 225). Print.
- . *Comunicazione della Commissione al Parlamento Europeo, al Consiglio e al Comitato economico e sociale europeo. Sull’informazione scientifica nell’era digitale: accesso, diffusione e conservazione comunicazione*. Bruxelles, 2007. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0056:FIN:IT:PDF>. (Cit. on pp. 216, 217). Print.
- . *Proposta di decisione del Consiglio che stabilisce il programma specifico recante attuazione del programma quadro di ricerca e innovazione (2014-2020) - Orizzonte 2020*. Bruxelles, 2011. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0811:FIN:it:PDF>. (Cit. on p. 220). Print.
- . *Raccomandazione della Commissione, del 17 luglio 2012, sull’accesso all’informazione scientifica e sulla sua conservazione*. Bruxelles, 2012. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:IT:PDF>. (Cit. on pp. 209, 217, 218, 220). Print.
- De Robbio, Antonella and Silvia Giacomazzi. “Dati aperti con LODE”. *Bibliotime* 14.2. (2011). (Cit. on p. 212). Web.

- Key Perspectives. *Data dimensions: disciplinary differences in research data sharing, reuse and long term viability. SCARP Synthesis Study*. Digital Curation Centre, 2007. <http://www.dcc.ac.uk/sites/default/files/documents/publications/SCARP%20SYNTHESIS.pdf>. (Cit. on p. 223). Print.
- Ministero dell'Istruzione, dell'Università e della Ricerca. *Horizon 2020 Italia*. 2013. <https://www.researchitaly.it/uploads/50/HIT2020.pdf>. (Cit. on pp. 225, 226). Print.
- Murray-Rust, Peter. "Open Data in Science". *Serials Review* 34.1. (2008): 52–64. (Cit. on pp. 212, 219). Web.
- National Science Foundation. National Science Board. *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*. Washington: National Science Board, 2005. [http://www.nsf.gov/nsb/meetings/2005/LLDDC\\_draftreport.pdf](http://www.nsf.gov/nsb/meetings/2005/LLDDC_draftreport.pdf). (Cit. on pp. 212, 214). Print.
- Organisation for Economic Cooperation and Development. *Principles and Guidelines for Access to Research Data from Public Funding*. Paris: OECD Publications, 2007. <http://www.oecd.org/science/scienceandtechnologypolicy/38500813.pdf>. (Cit. on pp. 210, 214, 221, 224). Print.
- Suber, Peter. "Promoting open access in the humanities". *Syllecta Classica* 16. (2005): 231–246. (Cit. on pp. 213, 219). Web.
- Swan, Alma. *Policy Guidelines For The Development And Promotion Of Open Access*. Paris: UNESCO, 2012. <http://unesdoc.unesco.org/images/0021/002158/215863e.pdf>. (Cit. on pp. 209, 222). Print.

MAURA FUNARI, laureata in Storia e Scienze della Documentazione,  
Università degli Studi di Perugia.

[funmaura@libero.it](mailto:funmaura@libero.it)

---

Funari, M. "Research data and humanities: a European context". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #8927. DOI: [10.4403/jlis.it-8927](https://doi.org/10.4403/jlis.it-8927). Web.

ABSTRACT: The article holds an inquiry on research data in the field of humanities, into the European zone. Defining these kind of data as primary source and basis of the scientific research's results, specificities characterizing them, in humanities, have been individualized. The attention paid to research data in European policies confirms their strategic role to the development and optimization of the scientific research. The analysis of a few research infrastructures and projects focused on research data in humanities, shows the state's policy central role to improve and develop them, making individual activities efficient and restricting a lack of homogeneity, supporting the presence of easily identifiable, usable and integrated systems. In this context, up until today, Italian rear position referring to systems of collection, management, preservation and sharing of the research data, seems to place itself in development's prospects.

KEYWORDS: Research data; Humanities; Research Infrastructures; Open Access; Open Data.

---

Submitted: 2013-06-04

Accepted: 2013-08-21

Published: 2014-01-01







# MOOC e biblioteche accademiche: tra possibilità di "esserci" e criticità

Laura Testoni

## 1 MOOC e biblioteche accademiche: tra possibilità di "esserci" e criticità

L'articolo prende in esame i Massive Open Online Courses (MOOC), una forma di editoria didattica accademica affermatasi nel 2012 soprattutto negli Stati Uniti. Nella parte iniziale dell'articolo vengono descritte le caratteristiche dei MOOC, la loro genesi, i possibili sviluppi e le criticità. Vengono poi discusse le modalità con le quali le biblioteche accademiche possono partecipare all'elaborazione dei MOOC, con particolare attenzione alle riflessioni poste dalla letteratura e da quanto elaborato da alcune organizzazioni come OCLC, American Library Association (ALA) e, in Europa, da Joint Information Systems Committee (JISC). Nella parte conclusiva si discutono le possibili interazioni tra biblioteche accademiche e MOOC alla luce del modello dell'*embedded librarianship*, e tenendo conto degli spunti offerti dall'*information literacy* e delle tematiche relative all'Accesso aperto.



## 2 Cosa sono i MOOC?

I MOOC sono corsi erogati in modalità esclusivamente online da istituzioni accademiche. Pur affondando le loro radici nella lunga tradizione della formazione a distanza, i MOOC sembrano rappresentare una autentica novità (Clarke) nel panorama accademico (soprattutto negli Stati Uniti), a causa del loro successo subitaneo e globale, testimoniato dalle iscrizioni in massa.

I principali aspetti che caratterizzano i MOOC sono:

- accesso gratuito previa registrazione;
- disponibilità online: chiunque abbia un accesso a Internet può iscriversi;
- partecipazione su larga scala: non ci sono limiti numerici alle iscrizioni; ne consegue che alcuni corsi sono simultaneamente fruiti da decine di migliaia di iscritti sparsi in tutto il mondo;
- scalabilità: è possibile fruire un solo video per pochi minuti o seguire tutto il percorso didattico svolgendo tutte le attività proposte;
- ubiquità dell'apprendimento: ogni MOOC presuppone che sia possibile apprendere ovunque e non necessariamente in una classe; non ci sono confini spazio-temporali all'educazione;
- feedback elettronico: i docenti (in genere un team) sono presenti nei video e in alcuni casi presidiano i forum online, ma il feedback è prevalentemente automatico;
- *branding*: i corsi hanno il marchio dell'università che li propone;
- il design permette e incoraggia il massimo di interattività tra i partecipanti;

- piena integrazione con i social network;
- conseguimento di un certificato finale che attesta l'acquisizione dei contenuti proposti attraverso il superamento di test o la consegna di elaborati.

### 3 I precursori

I MOOC si inseriscono nella storia della formazione a distanza. Come sottolineato da Nicholas Carr («The Crisis in Higher Education») gli anni '20 del secolo scorso furono negli Stati Uniti il grande periodo dei corsi per corrispondenza, promossi dallo *Home Study Department* della Chicago University attraverso due media universali: le poste e poi la radio<sup>1</sup>. Nel 1971 iniziò le sue attività nel Regno Unito la Open University<sup>2</sup>, che inizialmente erogava corsi a pagamento via posta, radio e televisione, ed è tutt'ora attiva con propri canali su iTunes\_U<sup>3</sup>, YouTube e una propria piattaforma MOOC<sup>4</sup>.

A partire dagli anni Zero sia iTunes\_U che YouTube sono stati (e sono tutt'ora) i canali attraverso i quali numerosi atenei rilasciano, con finalità prevalentemente promozionali, i propri materiali didattici. Il Massachusetts Institute of Technology (MIT), attraverso l'iniziativa *OpenCourseWare*, mette a disposizione gratuitamente i

---

<sup>1</sup>Si tratta dell'esperienza dei cosiddetti "College of the air": nel 1922 l'Università di New York aprì un canale radiofonico, seguita da moltissime università americane (Matt e Fernandez)

<sup>2</sup><http://www.open.ac.uk>. Tutti i link presenti nelle note e in bibliografia sono stati verificati l'8 settembre 2013.

<sup>3</sup>Attualmente (settembre 2013) iTunes\_U, il canale dello store Apple dedicato alle Università, dichiara di permettere l'accesso a 500.000 risorse didattiche gratuite <http://www.apple.com/it/education/itunes-u>.

<sup>4</sup>La piattaforma è Future learn <http://futurelearn.com> che ha iniziato i propri corsi a metà settembre 2013 in partenariato con numerose università inglesi e altre istituzioni culturali britanniche (British museum, British library e British Council)

suoi materiali didattici dal 2002. Ad oggi permette l'accesso a più di 2.100 corsi<sup>5</sup>, e fa a sua volta parte dell'iniziativa OpenCourseWare Consortium (OCWC), un gateway che permette l'accesso unificato a tutti i materiali didattici di differenti università e istituti di ricerca, rilasciati con licenza *Creative Commons attributions 3.0*<sup>6</sup>.

Nel 2008 Stephen Downes e George Siemens, due docenti dell'università canadese di Athabasca, specializzata in formazione a distanza, rilasciarono un corso aperto online che per la prima volta venne chiamato MOOC, a cui si registrarono 2.800 studenti. Alla base di questi corsi c'era uno specifico modello pedagogico, detto Connettivismo (Siemens) in base al quale la conoscenza non è un elemento statico e individuale ma è generata dalle connessioni tra persone e tra nodi informativi, e la vitalità dei nodi che connettono le reti di persone e contenuti è alla base dell'apprendimento permanente. Nello stesso anno l'università di Stanford lanciò l'iniziativa *Stanford Engineering Everywhere*<sup>7</sup> rilasciando online alcuni corsi, software, tutorial di computer science e intelligenza artificiale sia nel proprio sito che nei consueti canali iTunes\_U e YouTube.

## 4 L'affermazione dei MOOC: Udacity, Coursera, EdX

Il New York Times (Pappano) indica come "L'anno dei MOOC" il 2012, ed in effetti proprio in quel periodo i MOOC raggiunsero

---

<sup>5</sup><http://ocw.mit.edu/about/our-history>.

<sup>6</sup>La licenza Creative Commons 3.0 (CC BY-NC-SA 3.0 US) permette di copiare, distribuire, trasmettere i materiali didattici, di remixarli e riadattarli a condizione di citarne la fonte, di farne un uso non commerciale e di ridistribuire eventuali opere remixate con la medesima licenza. I contenuti estesi della licenza sono consultabili all'url: [http://creativecommons.org/licenses/by-nc-sa/3.0/us/deed.en\\_US](http://creativecommons.org/licenses/by-nc-sa/3.0/us/deed.en_US).

<sup>7</sup><http://see.stanford.edu>.

il picco di massima aspettativa, tipico del ciclo di hype descritto da Gartner<sup>8</sup>. Il successo è certamente collegato all'emergere di tre startup, fortemente legate ad istituzioni universitarie americane: Udacity, Coursera e EdX.

A gennaio 2012 Sebastian Thrun, docente di intelligenza artificiale a Stanford e membro dello staff Google, fonda Udacity, che inizialmente offre corsi di computer science. Oggi Udacity offre 28 corsi<sup>9</sup> rilasciati con licenza Creative Commons e, secondo l'Economist, conta 475.000 studenti<sup>10</sup>. Sempre nello stesso anno viene fondata dal MIT e dall'università di Harvard EdX, che attualmente dichiara 67 corsi,<sup>11</sup> 32 università partner e, secondo il Time (Ripley), 350.000 studenti.

Il 10 settembre 2013 EdX annuncia un partenariato con Google per la creazione di MOOC.org, una nuova piattaforma open source.<sup>12</sup> Nel momento in cui scriviamo la piattaforma MOOC.org è solo un annuncio e non una concreta realizzazione, e quindi per questo motivo ci pare prematuro leggere le implicazioni dell'ingresso di Google in questo scenario: ci sembra tuttavia chiaro che il progetto Google Libri può produrre sinergie con una piattaforma didattica a cui partecipano università e istituzioni di insegnamento superiore; inoltre, come reso esplicito nello stesso comunicato congiunto che

---

<sup>8</sup>Ciclo di hype descrive la "fortuna" nella ricezione delle tecnologie. All'interno del ciclo il "picco di aspettativa" corrisponde all'iper-esposizione del tema, che genera entusiasmo irrealistico ed ingiustificato. Cfr. <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp>.

<sup>9</sup>Cfr. <http://www.udacity.com/courses> che riporta la lista dei corsi: il dato si riferisce alla data di consultazione del sito (8 settembre 2013)

<sup>10</sup><http://www.economist.com/news/international/21568738-online-courses-are-transforming-higher-education-creating-new-opportunities-best>.

<sup>11</sup><https://www.edx.org/course-list/allschools/allsubjects/allcourses> che riporta la lista dei corsi: il dato si riferisce alla data di consultazione del link (8 settembre 2013)

<sup>12</sup>All'url <http://mooc.org/press.html> si trova il comunicato stampa congiunto che illustra a grandi linee il progetto

annuncia il partenariato, Google è interessata ad esplorare “come gli studenti apprendono”: gli stili cognitivi ci pare siano un aspetto cruciale nell’ambito della progettazione di tutti i servizi Google, e la gestione di una piattaforma MOOC ci pare rappresenti, in questo senso, un osservatorio privilegiato.

A maggio 2012 viene fondata Coursera, che ha come partner originario l’università di Princeton, ma collabora con 10 università americane e circa 75 università e istituzioni culturali nel mondo<sup>13</sup>. Attualmente (settembre 2013) Coursera dichiara<sup>14</sup> 4.640.000 studenti e 437 corsi, affermandosi come la realtà più consistente tra quelle sopra descritte.

Coursera e Udacity, a differenza di EdX, sono aziende for-profit, fondate e in parte finanziate dalle stesse università, che conferiscono a queste piattaforme – come terze parti – i loro corsi. Le piattaforme MOOC presentano i corsi di differenti università in un unico ambiente omogeneo. Nella letteratura consultata i MOOC sono spesso qualificati come una innovazione “dirompente” (*disruptive*) nella formazione superiore. *Disruptive* è qualunque processo attraverso il quale un settore che inizialmente serviva solo poche persone poiché i suoi prodotti e servizi erano complicati, dispendiosi e inaccessibili, si trasforma proponendo prodotti e servizi semplici, più convenienti e comodi che possono essere usati da molte persone, a prescindere dal loro reddito o dalla loro abilità (Christensen et al.).

Vale la pena di sottolineare la differenza tra i MOOC la mera disponibilità di materiali didattici in rete (modello OpenCourseWare): la maggior parte dei MOOC prevede ovviamente la possibilità di

---

<sup>13</sup>Cfr. la lista dei partner: <http://www.coursera.org/partners>, il dato è stato osservato l’8 settembre 2013. Tra i partner segnaliamo, unica realtà italiana, l’Università La Sapienza di Roma, che sta preparando tre corsi: <http://www.coursera.org/sapienza>.

<sup>14</sup>Cfr. <http://www.coursera.org/> la homepage di Coursera riporta il contatore con il numero degli iscritti, che cambia in tempo reale. Il contatore è stato osservato l’8 settembre 2013.

accedere ai materiali didattici (testi e video), ma punta ad un approccio esperienziale all'apprendimento: l'allievo può esercitarsi con domande a quiz che permettono di verificare il livello raggiunto ("check yourself"), tool didattici 3D<sup>15</sup>, tutorial, wiki, simulazioni, veri e propri compiti da svolgere o elaborati da realizzare e consegnare, esami finali. In alcuni casi il superamento di un MOOC permette di ottenere crediti formativi: l'ACE (*American Council on Education*) ha preso in esame la possibilità di associare crediti formativi al superamento di alcuni corsi.<sup>16</sup> Soprattutto, l'allievo può confrontarsi in modalità paritaria con gli altri studenti sparsi nel mondo attraverso sistemi di messaggistica interni al corso (forum) oppure su spazi dedicati in piattaforme di terze parti, come Meetup. Alcuni corsi integrano un profilo Facebook e Twitter come vero e proprio spazio didattico e di confronto tra i partecipanti.

## 5 Le ragioni del successo e le criticità

Al di là della moda del momento, e dall'enfasi che viene associata alle novità, il successo dei MOOC può essere spiegato da diversi fattori, di carattere economico e sociale.

Il costo dell'insegnamento superiore di qualità negli Stati Uniti è aumentato di 12 volte negli ultimi 30 anni<sup>17</sup> al punto che è molto diffusa la pratica di aprire linee di credito specifiche per le famiglie o

<sup>15</sup>Nelle discipline scientifiche ad esempio modelli di strutture molecolari.

<sup>16</sup>A novembre 2012 l'*American Council on Education* annuncia che cinque corsi di Coursera sarebbero stati valutati ammissibili per il conseguimento di crediti formativi, cfr. <http://www.acenet.edu/news-room/Pages/ACE-to-Assess-Potential-of-MOOCs-Evaluate-Courses-for-Credit-Worthiness.aspx>. A febbraio 2013 Coursera annuncia ufficialmente che cinque suoi corsi permettono di conseguire crediti: cfr. <http://blog.coursera.org/post/42486198362/five-courses-receive-college-credit-recommendations>.

<sup>17</sup><http://www.bloomberg.com/news/2012-08-15/cost-of-college-degree-in-u-s-soars-12-fold-chart-of-the-day.html>.

per gli studenti che accedono all'educazione superiore. Secondo uno studio recente del PEW Centre (*Is College Worth It? College Presidents, Public Assess, Value, Quality and Mission of Higher Education*), il costo medio per la frequenza al college è attualmente di 27.000 dollari americani all'anno (e arriva a 50.000 nelle strutture più elitarie). Dallo studio emerge che, sebbene per il 75% degli intervistati le spese per l'istruzione superiore siano troppo onerose, più del 94% vorrebbe che il proprio figlio potesse accedere a livelli elevati di istruzione.

Sebbene i costi dell'istruzione superiore negli Stati Uniti siano molto elevati non è solo questo il principale fattore esplicativo del successo dei MOOC. Scientific American (Waldrop) rileva che gli utenti di Coursera appartengono a 220 differenti nazionalità (tra cui spiccano i Paesi emergenti) mentre gli americani rappresentano il 27% dell'utenza. La crescente diffusione di internet a livello globale, l'affermarsi nei Paesi emergenti di sempre più elevati livelli di scolarizzazione<sup>18</sup>, e quindi di fabbisogni formativi crescenti e sofisticati, sono certamente fattori che rendono i MOOC un'esperienza su cui converge l'attenzione, e comunque un progetto coerente con una società aperta e globalizzata. Il successo dei MOOC nei Paesi emergenti segnala inoltre che, in quei Paesi, oltre all'alfabetizzazione primaria, cavallo di battaglia delle tradizionali Organizzazioni Non Governative e agenzie educative, si fa sempre più pressante la

---

<sup>18</sup>Cfr. UNESCO Institute for Statistics. I dati relativi all'accesso all'educazione terziaria dal 2000 al 2010 aumentano rispettivamente del 135,59% in Brasile, del 119,6% in India e del 321,5% in Cina. Per avere un'idea degli ordini di grandezza segnaliamo che nel 2010 sono rispettivamente più di 6 milioni, più di 20 milioni e più di 31 milioni le persone che accedono all'istruzione superiore rispettivamente in Brasile, India e Cina. I valori fanno riferimento a tutti i generi (maschi+femmine). Abbiamo ricavato questi dati utilizzando Undata, il portale statistico delle Nazioni Unite. Il link permanente alla tabella generata dalla nostra ricerca è il seguente: <http://goo.gl/ftZFbZ>.

necessità di un'educazione superiore di carattere tecnico scientifico, e di alto livello.

Non a caso, le piattaforme che offrono MOOC considerano un punto d'onore avere utenti iscritti da molti Paesi nel mondo. Un'utenza globalizzata conferma infatti la "*vision*" che sembra animare l'insieme dei corsi: la diffusione nel mondo di formazione specialistica e di alto livello con modalità libere e senza frontiere.

I MOOC, che sono stati chiamati "*Ivy League per le masse*" (Ripley), permettono l'accesso a materiali rilasciati con il marchio di università prestigiose e offrono un'esperienza didattica di diverse settimane, rilasciando, in alcuni casi dietro il pagamento di una piccola quota, certificati individuali che ovviamente non hanno il valore di un vero e proprio diploma, ma che comunque attestano il superamento di prove scritte e il conseguimento di competenze.

Vale la pena di passare in rassegna le maggiori criticità legate al fenomeno dei MOOC.

**Quale business model?** Il modello di business (e quindi la sostenibilità a lungo termine) di questi corsi non è ancora del tutto evidente (Korn e Levitz): è chiaro che per le università che "firmano" i corsi e impegnano i loro docenti più prestigiosi, il MOOC è uno strumento per promuovere l'offerta didattica, e attirare così, in teoria, nuovi studenti. Inoltre i docenti offrono gratuitamente libri di testo e appunti come materiali necessari e sufficienti per l'apprendimento, ma possono segnalare le loro pubblicazioni, come lettura opzionale, agli studenti iscritti. Non sfugge che gli editori accademici possano vedere nell'utenza potenzialmente illimitata dei MOOC un mercato interessante (Howard).<sup>19</sup>

---

<sup>19</sup> Anche i grandi editori stanno negoziando accordi con le piattaforme MOOC a cui offrono, a libero accesso, libri di testo in modalità streaming, nella speranza che i partecipanti acquistino la più maneggevole versione ePub o cartacea (Kolowich).

**Quale didattica?** Proprio per il loro carattere "massivo" (alcuni corsi possono avere decine di migliaia di iscritti alla volta) nei MOOC non è in genere<sup>20</sup> previsto il rapporto diretto con il docente presente nei video didattici. La correzione dei compiti è automatica (quiz, test a risposta multipla gestiti dal software) oppure avviene tra pari in modalità *crowdsourcing*<sup>21</sup>: il compito è sottoposto a una lista casuale di partecipanti, che danno un voto, e il voto finale è la media dei voti ottenuti. Questo tipo di validazione dei compiti, non problematica per le discipline tecnico-scientifiche, può essere più delicata per le scienze sociali o umane<sup>22</sup>.

**I drop-out:** un altro aspetto rilevato da più fonti (Clarke; Morris; Parr; Webb) è l'alto tasso di abbandono: meno del 10% di coloro che si iscrivono a un MOOC lo portano a termine, e quindi il numero elevatissimo di "iscritti" non corrisponde al numero di coloro che completano il corso: la maggioranza, anzi, si iscrive "per provare" e abbandona a un certo punto il percorso didattico.

**Il modello pedagogico:** i primi MOOC rilasciati nel 2008 si basavano, come si è visto, sul modello connettivista, che presuppone una vera e profonda interazione reticolare tra gli individui. È stato osservato (*NMC horizon report: 2013 Higher Education Edition*) che i MOOC seguiti simultaneamente da centinaia di

---

<sup>20</sup>Alcuni corsi prevedono un "office hour" (cioè una interlocuzione diretta e programmata con i formatori, che ricalca il modello del colloquio nello studio del docente) interno al MOOC, oppure su piattaforme terze come Google Hangout.

<sup>21</sup>Si tratta del così detto "peer grading": "Every student's work is evaluated by multiple students based on a predefined grading rubric and then, the average of multiple scores is used as a grade for that work" (Subbian).

<sup>22</sup>Non riteniamo sia un caso che i cinque corsi ammessi dall'American Council on Education (ACE) per i crediti formativi riguardino *solo* discipline STEM (science, technology, engineering, and mathematics).

migliaia di persone non possono sostenere questo modello, e sono basati prevalentemente su modalità più povere, che prevedono il rilascio di video e la verifica di quanto appreso attraverso assegnazione di quiz (corretti dal software) e compiti: l'attività creativa di gruppo, sebbene enfatizzata, resta sullo sfondo.

## 6 MOOC e biblioteche accademiche: una sfida, alcuni problemi

I MOOC fanno ormai pienamente parte, almeno negli Stati Uniti, del panorama accademico.

Proprio per questo le biblioteche accademiche hanno preso in considerazione questo fenomeno per capire se e come possono, dalla loro posizione, partecipare all'elaborazione di questa nuova ondata di contenuti formativi di massa<sup>23</sup>. Le difficoltà non mancano (Wright): il carattere massivo dei MOOC rende difficile per le biblioteche applicare gli approcci abituali della formazione a distanza "a numero chiuso", che comprendono di solito la proposta di navigazione nelle piattaforme dei sistemi bibliotecari di Ateneo, la ricerca di documenti in cataloghi e discovery tool ecc. Molti MOOC sono ospitati su piattaforme terze (ad esempio, come si è visto, Coursera, EdX, Udacity) in cui i bibliotecari non possono attivare modalità di interfacciamento diretto con la biblioteca (esempio: pulsante "Chiedi al bibliotecario"), possibili invece se si opera nella piattaforma e-learning dell'istituzione in cui si opera. Ciò significa che le biblioteche dovrebbero creare unità didattiche "*plug and play*", cioè

---

<sup>23</sup>Sono all'incirca questi gli interrogativi presenti in un white paper dell'ACRL (Association of College and Research Libraries, l'associazione delle biblioteche di ricerca americane): (Butler)

strutturate in modo da essere immediatamente applicabili a qualunque piattaforma didattica: la metafora è quella dei dispositivi hardware "*plug and play*", che non richiedono adattamenti e configurazioni specifiche, ma iniziano a funzionare immediatamente non appena collegati ad un computer, qualunque esso sia. La maggior parte delle risorse informative offerte dalle biblioteche accademiche prevedono licenze campus che non sono applicabili ai MOOC, per definizione aperti a chiunque abbia un accesso a Internet. È stato inoltre osservato (Wu) che il modello di business dei MOOC può evolversi verso una "*iTunization*" dell'informazione: gli utenti dei corsi acquistano testi e materiali didattici in modo granulare a costi molto bassi (come un brano su iTunes) direttamente dagli editori partner dei MOOC, lasciando alle biblioteche accademiche il ruolo di "club esclusivo" per gli studenti che possono permettersi l'esperienza in un vero campus.

## 7 Quali ambiti di intervento per le biblioteche accademiche?

In premessa si ritiene sarebbe utile, per i bibliotecari, iscriversi a uno o più MOOC, per vedere come funzionano. Capirne i meccanismi didattici, coglierne dal vivo il ritmo, le criticità e i punti di forza è comunque un buon approccio per assumere quella "*MOOC fluency*" (Matthews) che può permettere ulteriori e più incisivi interventi. Si possono scegliere MOOC che corrispondono ai propri interessi personali; ma un *subject librarian* può seguire MOOC riguardanti l'ambito disciplinare di interesse, e esistono comunque dei MOOC di biblioteconomia.<sup>24</sup>

---

<sup>24</sup>Segnaliamo "New librarianship master class", un MOOC organizzato dalla School of information studies della Syracuse university. Il corso, iniziato a luglio 2013 si è concluso mentre scriviamo (settembre). Gestito da David Lankes

Ciò premesso gli ambiti di intervento per le biblioteche accademiche potrebbero essere alcuni tra i seguenti.

**Collaborare alla realizzazione di un MOOC.** Come si è visto i MOOC sono contenuti didattici complessi che includono video, testi, esercitazioni: la loro progettazione e realizzazione non viene quindi svolta da un solo docente (esperto della materia oggetto del corso) ma richiede un lavoro di staff. Il bibliotecario che collabora alla realizzazione di un MOOC esce dalle mura (reali e virtuali) della biblioteca e offre consulenza in attività su cui detiene competenze specifiche.

- Il reference librarian identifica e localizza le risorse appropriate per le necessità didattiche dei MOOC (testi, immagini, video)(Massis). Per "appropriate" si intende non solo precise, autorevoli, attuali e obiettive, ma anche utilizzabili nel contesto di massima apertura tipico dei MOOC, che rende necessario l'utilizzo di risorse di pubblico dominio o contrassegnate da licenze aperte (Becker)
- Un'attività essenziale consiste nel supportare e consigliare nella gestione delle tematiche legali connesse al copyright che sorgono dall'operare in un ambiente online completamente aperto: ad esempio negli Stati Uniti nel

---

e da suoi collaboratori approfondisce i temi affrontati nell'*Atlas of new librarianship* <http://ischool.syr.edu/future/grad/newlibopencourse.aspx>. Ma in senso lato non possiamo non considerare corsi interessanti per un bibliotecario: "Economie du document web", MOOC in francese promosso tra gli altri dall'ENSSIB <http://archinfo24.hypotheses.org/>; "Understanding media by understanding Google" della Northwestern University <http://www.coursera.org/course/googlemedia>, oppure: "Metadata: organizing and discovering information" dell'University of North Carolina at Chapel Hill <http://www.coursera.org/course/metadata>.

contesto dei MOOC il *fair use* dovrebbe essere la regola<sup>25</sup>, anche se il livello di visibilità e l'utenza tendenzialmente illimitata e universale dei corsi sembrano richiedere una maggiore cautela.

Segnaliamo, come buone prassi, le linee guida per i corsi online elaborate dal sistema bibliotecario dell'Università di Stanford<sup>26</sup> e il *MOOC copyright service* organizzato dalle biblioteche della Duke University per facilitare il corretto uso dei materiali didattici (Fowler). Va rilevato che in questo ambito alcuni sistemi bibliotecari utilizzano i servizi di agenzie terze, specializzate nel gestire le tematiche connesse al copyright<sup>27</sup>.

**Svolgere attività di advocacy e promozione** delle biblioteche e dell'accesso aperto:

- Le biblioteche, secondo l'Association of Research Libraries<sup>28</sup> (Butler) possono fare pressione affinché i MOOC, che sono essenzialmente un tipo molto particolare di editoria accademica, optino per l'utilizzo di materiali didattici aperti: in effetti i MOOC sono *free* nel senso che sono gratuiti e ad accesso libero (previa registrazione); tuttavia il regime dei materiali offerti dalle grandi piattaforme

---

<sup>25</sup>Butler presenta una disamina degli aspetti problematici che legano il *fair use* didattico, codificato dal *Copyright Act*, in relazione alla specificità dei MOOC, che hanno una utenza potenzialmente illimitata e globale (*Massive Open Online Courses: Legal and Policy Issues for Research Libraries*). Segnaliamo che ARL (Association of Research Libraries) ha elaborato un *Code of best practices in fair use for academic and research libraries* <http://www.arl.org/focus-areas/copyright-ip/fair-use/code-of-best-practices>.

<sup>26</sup>Stanford online course guidelines <https://library.stanford.edu/using/copyright-reminder/stanford-online-course-guidelines>.

<sup>27</sup>Ad esempio il sistema bibliotecario dell'Università di Stanford, fondatrice della piattaforma Edx, utilizza i servizi dell'agenzia SIPX <http://www.sipx.com> per gestire i problemi di copyright emergenti dai materiali inclusi nei MOOC (Calter).

<sup>28</sup>ARL, l'associazione delle biblioteche di ricerca americane.

(Coursera EdX Udacity), nonché il MOOC come prodotto didattico unitario, non presentano licenze aperte, come emerge dall'analisi delle rispettive clausole di servizio (Kolowich).

- I bibliotecari possono collaborare con i docenti della propria istituzione impegnati nei MOOC segnalando loro la sezione "tutorial" della biblioteca, a condizione che i tutorial includano solo risorse liberamente accessibili, dal momento che, come si è visto, l'apertura potenzialmente illimitata dei MOOC non permette di utilizzare risorse vincolate da sottoscrizione, a meno che non sia stata stipulata apposita licenza d'uso.

**Produrre MOOC basati sui temi dell'*information literacy*.** Come abbiamo visto i MOOC registrano un livello molto elevato di abbandoni. Va evidenziato che utilizzare un MOOC con successo richiede la capacità di estrarre informazioni dai video didattici proposti, elaborare testi, interagire con altri in ambiente digitale. Si tratta di abilità e sensibilità vicine a quella che è stata definita *transliteracy*, una specifica modalità di *information literacy*. La *transliteracy* si identifica con ciò che significa essere *literate* nel 21° secolo, cioè nell'abilità di leggere, scrivere e interagire in modo fluido attraverso un insieme vasto di piattaforme, strumenti e media, ma anche nella capacità a relazionarsi con altri in un ambiente digitale. I bibliotecari sono esperti nel produrre materiale formativo, e potrebbero quindi produrre MOOC basati sulle competenze e attitudini proprie della *transliteracy* (Thomas): far adottare questi materiali potrebbe essere un primo e semplice passo per arricchire i MOOC di contenuti. Vale la pena segnalare, ad esempio, il MOOC "Information literacy for art and design students", elaborato da AICAD (Association of independent colleges of art and

design), un consorzio non-profit di 43 istituti di educazione superiore artistica americani e canadesi.<sup>29</sup>

**Studiare le possibilità per la conservazione e la digital curation dei materiali prodotti dai corsi** (progetti degli allievi, conversazioni generate dai forum). Questa attività presuppone tuttavia una maggiore chiarezza rispetto alla effettiva proprietà dei materiali didattici prodotti e generati dai MOOC (appartengono alle Università che li producono o alle piattaforme che li ospitano?).

## 8 Le iniziative internazionali: OCLC, ALA, JISC, Unione europea

Le grandi istituzioni bibliotecarie hanno iniziato ad interrogarsi sulla portata e l'impatto dei MOOC nelle loro attività.

Il 18 marzo 2013 OCLC ha organizzato un seminario<sup>30</sup> da cui emergono molte indicazioni pratiche per mettere a regime le potenzialità dei MOOC per le biblioteche accademiche; in questa occasione è stato attivato un gruppo di discussione permanente e aperto a tutti, per permettere di continuare il confronto.<sup>31</sup> L'ALA ha istituito

---

<sup>29</sup>Questo MOOC ha per tema le strategie di ricerca e le tecniche di fil-traggio delle informazioni, e dichiara di "incoraggiare gli studenti a esplora-re le risorse trovate nelle biblioteche locali", <https://www.canvas.net/courses/information-literacy-for-art-and-design-students>.

<sup>30</sup>"MOOCs and libraries: massive opportunity or overwhelming challenge?", <http://www.oclc.org/research/events/2013/03-18.html>.

<sup>31</sup>Il gruppo, fondato a marzo 2013 si intitola "MOOCs and Librarianship" e, dopo una parentesi iniziale su Google group (<https://groups.google.com/forum/#!forum/moocs-and-librarianship>) è attualmente ospitato presso i server dell'ALA, e è diven-tato strumento di comunicazione di uno specifico gruppo di lavoro istituito in seno all'American Library Association (ALA), l'associazione dei bibliotecari americani, <http://lists.ala.org/sympa/info/acr-dgmoocs>.

nel corso del suo congresso annuale del 2013, un gruppo di di lavoro sul supporto ai MOOC,<sup>32</sup> con l'obiettivo di esplorare il ruolo delle biblioteche per quanto concerne il copyright, il *fair use*, le licenze e il supporto a studenti e staff.

In Europa JISC, (Joint Information Systems Committee, il consorzio che offre servizi e tecnologia alle biblioteche di ricerca del Regno Unito), ha organizzato un webinar sui MOOC a luglio 2013,<sup>33</sup> con un approccio più problematico rispetto ad OCLC: ciò che maggiormente sembra oggetto di perplessità è il modello "industriale" americano dei MOOC da decine di migliaia di studenti, in cui il brand (il nome dell'università prestigiosa) sembra prevalere sui contenuti o sugli aspetti pedagogici.

La posizione espressa da JISC testimonia come in Europa i MOOC non abbiano ancora capitalizzato quell'interesse tributato loro negli Stati Uniti:<sup>34</sup> il sistema accademico europeo è differente, e forse meno orientato al marketing e alla serrata competizione tra atenei; la formazione a distanza, a parte l'eccezione della Open University di Londra (che come s'è visto ha aperto a metà settembre una propria piattaforma MOOC), non è così radicata; persistono (sempre più tenui) le barriere linguistiche.<sup>35</sup>

Va comunque rilevato che la Commissione europea il 25 settembre 2013 lancia il piano di azione *Opening up education* con l'intenzione di incentivare l'innovazione e le competenze digitali nelle scuole

---

<sup>32</sup>The library support for massive open online courses (MOOCs) discussion group, <http://www.ala.org/acrl/aboutacrl/directoryofleadership/discussiongroups/acr-dgmoocs>.

<sup>33</sup>"What is a MOOC? Webinar", <http://www.jisc.ac.uk/events/what-is-a-mooc-11-jul-2012>.

<sup>34</sup>Rammentiamo però che diverse università europee, tra cui La Sapienza di Roma, hanno inserito o stanno inserendo loro corsi nella piattaforma Coursera

<sup>35</sup>Non abbiamo sottolineato che l'ovvia koiné dei MOOC è l'inglese. Su Coursera, che come s'è visto presenta una elevata varietà di corsi, rileviamo che su 435 corsi 402 sono in inglese (rilevazione effettuata l'8 settembre 2013).

e nelle università (Unione Europea). Nel documento che descrive questo piano di azione i MOOC, insieme alle altre risorse didattiche aperte, sono considerati un'opportunità per aprire maggiormente i contesti di apprendimento. Il documento propone l'adozione di alcune linee strategiche per incoraggiare a sperimentare nuovi modelli didattici attraverso l'utilizzo di fondi strutturali e programmi come "Erasmus+" e "Orizzonte 2020".

Il documento, ampio e articolato, si sofferma anche sulla necessità di ricorrere a licenze aperte per la produzione di materiali didattici e sull'opportunità di adottare software aperti e piattaforme interoperabili, che non compromettano la possibilità, per ogni studente, di utilizzare il proprio dispositivo per la fruizione dei contenuti. La prima concreta tappa del progetto è rappresentata dall'apertura di un "European MOOC scoreboard" un portale che aggrega tutti i MOOC prodotti in Europa, proponendosi come gateway e punto di accesso unitario alle differenti risorse.

## 9 Conclusioni

I MOOC alludono a una dimensione globalizzata e non convenzionale della formazione di livello accademico. Ci pare possano rappresentare una sfida e un elemento di interesse per le biblioteche accademiche per diverse motivazioni.

Un MOOC è un prodotto complesso la cui elaborazione richiede competenze differenti, che eccedono quelle strettamente scientifiche del docente che eroga il corso: è necessario infatti produrre video, reperire e organizzare i materiali didattici, gestire le prove online. Un MOOC è il risultato di un lavoro di équipe.

Si ritiene quindi che i MOOC rappresentino per i bibliotecari accademici una opportunità per rafforzare i legami con la comunità scientifica, e per mettere alla prova, in modalità concrete, il model-

lo dell'*embedded librarianship* (Kesselman e Watsein).<sup>36</sup> Intendiamo con questo termine le attività professionali di carattere LIS (Library and information sciences) che vengono svolte fuori dal contesto abituale (la biblioteca), in unità organizzative composte da non-bibliotecari, all'interno delle quali il bibliotecario fornisce risorse e servizi informativi.

I MOOC, corsi universitari potenzialmente aperti a tutti, mostrano in modo evidente i limiti delle licenze tradizionali applicate alla letteratura scientifica, appropriate per una istituzione o per consorzi, ma inadatte a contesti didattici globali.

Per questa ragione i MOOC possono rappresentare un'opportunità per il movimento Open Access, che può essere il partner naturale di questo tipo di formazione accademica. Pensiamo non solo alla possibilità di utilizzare i repository istituzionali, ma anche piattaforme come DOAJ, DOAB, e alla possibilità di far emergere, all'interno di un MOOC, le pubblicazioni delle University press che pubblicano in regime OA, e che talvolta faticano a ottenere visibilità al di fuori dal proprio contesto accademico territoriale. Le biblioteche possono incoraggiare i docenti impegnati nei MOOC a pubblicare dispense e materiali didattici in prima battuta nelle piattaforme OA, linkando poi queste ultime ai MOOC, piuttosto che incorporare i loro materiali didattici direttamente dentro i MOOC, che sono, come s'è visto, *free* ma non necessariamente *open*.

Nel rapporto tra MOOC e information literacy, i MOOC rappresentano un banco di prova per la molteplicità di tutorial, guide, learning object prodotti dalle biblioteche accademiche nell'ambito delle attività di Information literacy: ma lo sforzo che andrà fatto sarà ripensare questi materiali in un ambiente non necessariamente

---

<sup>36</sup>Nello specifico l'articolo, pubblicato prima dell'emergere dei MOOC, sottolinea che la crescita dei materiali didattici online offre al bibliotecario l'opportunità di essere incorporato nell'elaborazione dei corsi, fornendo il supporto informativo e l'assistenza nel recupero dei documenti.

collegato alla biblioteca in cui si opera, adeguandoli a un contesto globale e senza confini.

## Riferimenti bibliografici

- Becker, Bernd W. «Connecting MOOCs and Library Services». *Behavioral & Social Sciences Librarian* 32. DOI: [10.1080/01639269.2013.787383](https://doi.org/10.1080/01639269.2013.787383). (2 2013): 135–138. (Cit. a p. 253).
- Butler, Brandon. *Massive Open Online Courses: Legal and Policy Issues for Research Libraries*. 2012. 3–13. <http://www.arl.org/storage/documents/publications/issuebrief-mooc-22oct12.pdf>. (Cit. alle pp. 251, 254).
- Calter, Mariellen. «MOOCs and the library: engaging with evolving pedagogy». *IFLA World Library and Information Congress, Singapore 17 - 23 August 2013*. (Cit. a p. 254).
- Carr, Nicholas. «The Crisis in Higher Education». *Technology Review* 115. (6 2013): 32–40. <<http://www.technologyreview.com/featuredstory/429376/the-crisis-in-higher-education/>>. (Cit. a p. 243).
- Christensen, Clayton, et al. «Center for American Progress. Disrupting college: how disruptive innovation can deliver quality and affordability to postsecondary education Center for American Progress». (2011). <[http://www.americanprogress.org/wp-content/uploads/issues/2011/02/pdf/disrupting\\_college.pdf](http://www.americanprogress.org/wp-content/uploads/issues/2011/02/pdf/disrupting_college.pdf)>. (Cit. a p. 246).
- Clarke, Thomas. «The advance of the MOOCs (massive open online courses). The impending globalization of business education?» *Education + Training* 55.4/5. DOI: [10.1108/00400911311326036](https://doi.org/10.1108/00400911311326036). (2013). (Cit. alle pp. 242, 250).
- Fowler, Lauren. «Drawing the blueprint as we build: setting up a library-based copyright and permissions service for moocs». *D-Lib Magazine* 19. DOI: [10.1045/july2013-fowler](https://doi.org/10.1045/july2013-fowler). (7/8 2013). (Cit. a p. 254).
- Howard, Jennifer. «Publishers see online mega-courses as opportunity to sell textbooks». *Chronicle of higher education*. (2012). <<http://chronicle.com/article/Can-MOOCs-Help-Sell/134446/>>. (Cit. a p. 249).
- Is College Worth It? College Presidents, Public Assess, Value, Quality and Mission of Higher Education*. Washington DC: Pew research Centre, 2011. 3–13. <http://www.pewsocialtrends.org/files/2011/05/higher-ed-report.pdf>. (Cit. a p. 248).
- Kesselman, Martin A. e Sarah Barbara Watsein. «Creating opportunities: embedded librarians». *Journal of Library Administration* 49. DOI: [10.1080/01930820902832538](https://doi.org/10.1080/01930820902832538). (4 2009): 383–400. (Cit. a p. 259).

- Kolowich, Steve. «Marketing to the MOOC masses». *Inside higher education*. (September 2012). <<http://www.insidehighered.com/news/2012/09/26/elsevier-partners-edx-provide-free-versions-textbooks-mooc-students>>. (Cit. alle pp. 249, 255).
- Korn, Melissa e Jennifer Levitz. «Online courses look for a business model. Free classes, open to the masses, seek to generate revenue from content licensing, exams or job-referral services». *Wall Street Journal January*. (2013). <<http://online.wsj.com/article/SB10001424127887324339204578173421673664106.html>>. (Cit. a p. 249).
- Massis, Bruce E. «MOOCs and the libraries». *New library world* 114. DOI: [10.1007/s10755-013-9263-2](https://doi.org/10.1007/s10755-013-9263-2). (5/6 2013): 267–270. (Cit. a p. 253).
- Matt, Susan e Luke Fernandez. «Before MOOCs, 'Colleges of the Air'». *Chronicle of Higher Education*. (2013). <<http://chronicle.com/blogs/conversation/2013/04/23/before-moocs-colleges-of-the-air/>>. (Cit. a p. 243).
- Matthews, Brian. «MOOC Fluency - some advice for future librarians». *The ubiquitous librarian*. (2012). <<http://chronicle.com/blognetwork/theubiquitouslibrarian/2012/06/27/mooc-fluency-some-advice-for-future-librarians/>>. (Cit. a p. 252).
- Morris, Libby V. «MOOCs, Emerging Technologies, and Quality». *Innovative Higher Education* 38. DOI: [10.1007/s10755-013-9263-2](https://doi.org/10.1007/s10755-013-9263-2). (4 2013): 251–252. (Cit. a p. 250).
- NMC horizon report: 2013 Higher Education Edition. Austin, TX: New Media Consortium, 2013. <http://www.nmc.org/pdf/2013-horizon-report-HE.pdf>. (Cit. a p. 250).
- Pappano, Laura. «The year of the MOOCs». *New York Times*. (2012). <<http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.htm>>. (Cit. a p. 244).
- Parr, Chros. «Mooc completion rates 'below 7%'. Open online courses' cohort much less massive at finish line». *Times of higher education*. (2012). <<http://www.timeshighereducation.co.uk/news/mooc-completion-rates-below-7/2003710.article>>. (Cit. a p. 250).
- Ripley, Amanda. «College is dead. Long live College! Can a new breed of online megacourses finally offer a college education to more people for less money?» *TIME*. (2012). <<http://nation.time.com/2012/10/18/college-is-dead-long-live-college/>>. (Cit. alle pp. 245, 249).
- Siemens, George. «What is the theory that underpins our moocs». *Elearnspace*. (2012). <<http://www.elearnspace.org/blog/2012/06/03/what-is-the-theory-that-underpins-our-moocs/>>. (Cit. a p. 244).
- Subbian, Vignesh. «Role of MOOCs in integrated STEM education: A learning perspective». *Integrated STEM Education Conference*. 2013. (Cit. a p. 250).

- Thomas, Sue. «Transliteracy: crossing divides». *First Monday*. (2007). <<http://journals.uic.edu/ojs/index.php/fm/article/view/2060/1908>>. (Cit. a p. 255).
- Unione Europea. *Aprire l'istruzione: tecniche innovative di insegnamento e di apprendimento per tutti grazie alle nuove tecnologie e alle risorse didattiche aperte. Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. Documento COM 654 final 25/09/2013*. 2013. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2013:0654:FIN:IT:PDF>. (Cit. a p. 258).
- Waldrop, Mitchell M. «Massive Open Online Courses, aka MOOCs, Transform Higher Education and Science». *Scientific American*. (2013). <<http://www.scientificamerican.com/article.cfm?id=massive-open-online-courses-transform-higher-education-and-science>>. (Cit. a p. 248).
- Webb, Jenn. «Visualization of the Week: MOOC completion rates». *O'Reilly Strata. Making data work*. (2013). <<http://strata.oreilly.com/2013/03/visualization-of-the-week-mooc-completion-rates.html>>. (Cit. a p. 250).
- Wright, Forrest. «What do librarians need to know about MOOCs?» *D-Lib Magazine* 19. (3/4 2013). DOI: [10.1045/march2013-wright](https://doi.org/10.1045/march2013-wright). (Cit. a p. 251).
- Wu, Kerry. «Academic Libraries in the Age of MOOCs». *Reference Services Review* 41. DOI: [10.1108/RSR-03-2013-0015](https://doi.org/10.1108/RSR-03-2013-0015). (3 2013). (Cit. a p. 252).

LAURA TESTONI, Biblioteca della Scuola di Scienze Sociali. Università di Genova.

[laura.test@gmail.com](mailto:laura.test@gmail.com)

---

Testoni, L. "MOOC e biblioteche accademiche". *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #9072. DOI: [10.4403/jlis.it-9072](https://doi.org/10.4403/jlis.it-9072). Web.

**ABSTRACT:** Massive open online courses(MOOCs) are online courses free and open to anyone: the MOOCs enrollement is essentially unlimited. Purpose of this essay is to offer an overview on the debate about the development of the MOOCs and about their impact on academic libraries information services. In the first part MOOCs are examined as an emerging educational tool and a very "disruptive" innovation in the field of the higher education. The development of MOOCs platforms as Coursera, EdX and Udacity in 2012 starts an enthusiastic phase (the so-called "peak of inflated expectations" in the Hype cycle model), followed by a wide debate in the educational / academic milieu. We consider the strengths and the critical points of MOOCs. In the second part the author focuses the attention on the interaction between the MOOCs and the academic libraries. The following issues are taken into account: the involvement / the roles of the librarians in the implementation of the MOOCs; the copyright management in the MOOCs environment; the information literacy; the library advocacy; the impact of the MOOCs on information services provided by academic libraries.

**KEYWORDS:** MOOC; e-learning; information literacy; lifelong learning; community; social networks; academic libraries

---

Submission: 2013-09-13  
Accettazione: 2013-10-27  
Pubblicazione: 2014-01-01







# MARC Formats in China: Local or International?

Ben Gu

## 1 Introduction

The use of MARC tapes was the first step for Chinese librarians to use MARC records. It can be traced back to 1980, when National Library of China (NLC) imported LC MARC tapes for test retrieval of books in Western languages. In July 1981, the Party School of the Central Committee of the Communist Party of China (CPC) imported LC MARC tapes containing 413,000 records during April 1979-April 1981. The tapes were used to retrieve books on international communist movement, and there were about 1000 hit records. Because Chinese libraries never used such new technology before, it should even be approved by the CPC Secretariat, the top level management at that time (An, “Yin jin MARC ci dai de shou ci ying yong”). The test work was done by MARC Cooperation Group jointly participated by the National Library of China, Peking University Library, Tsinghua University Library, Renmin University Library and the China National Publications Import & Export Corporation (“MARC ci dai yu wo guo xi wen bian mu”; Chen, “Beijing di qu xi wen tu shu ji du mu lu yan zhi jin zhan”). Actually, Chinese librarians began to pay attention to LC MARC as early as 1975, but they could not do anything but research, because the lack



of international exchange during the “Cultural Revolution” period (1966-1976) (Liu).

In 1985 Chinese libraries began to create bibliographical records of monographs in Chinese languages, and had accumulated about 1 million records by the end of 1990.

## 2 Decision of CNMARC Based on UNIMARC

Since the first publication of UNIMARC in 1977 and the second edition in 1980, Chinese library cataloging community, which was just considering catching up with international developments, decided to develop a Chinese local format based on UNIMARC, because it is an IFLA-developed international standard and even Library of Congress announced to implement UNIMARC (This what Chinese librarians knew about LC’s decision; actually, LC’s implementation was to provide conversion between USMARC and UNIMARC)(IFLA). There were other possible reasons for Chinese libraries to use UNIMARC as the basis of CNMARC:

- Main Entry Issue: Chinese libraries do not use main entry in the cataloging of Chinese materials, therefore USMARC, which was suitable for the cataloging with the concept of main entry, seemed strange to Chinese catalogers.
- MARC Expert Suggestion: It is said that Henriette Avram, the first developer of MARC format, visited China in late 1980s and expressed the feasibility for Chinese libraries to use UNIMARC.
- Japanese Example: Japanese libraries were already using MARC format earlier and basing its JapanMARC on UNIMARC.

- USMARC Shortcomings: There are some shortcomings with USMARC, and even other countries using USMARC made modifications and had their own versions.

Because there is a lack of historical records, the author can just provide the above reasons from the interviews with Mr. Yan ZHU, Ms. Beixin SUN and Mr. Yuanzheng CHEN, who were members participating in drafting of the first edition of CNMARC. People can hardly imagine the conditions of the 1970s and 1980s, when international exchange of information was extremely insufficient.

### 3 CNMARC and CMARC

The process of compiling the Chinese national format was the following: the National Library of China began to compile *Chinese MARC Communication Format* (later called CNMARC) in 1986, it completed a draft by the end of 1986 (Sun), sent copies to other libraries for national review in January 1987, and revised the draft according to the latest edition of UNIMARC (1987) in early 1988; the format was second-time reviewed during a special workshop hosted by the Library Society of China in 1989 (“Zhongguo ji du mu lu ge shi xue shu yan tao hui ji yao”), and was finally published by the Bibliography and Documentation Publishing House in February 1991. At the same time, the National Library began to distribute CNMARC records to libraries in China and around the world (“Beijing tu shu guan xiang guo nei wai fa xing ji du mu lu”). More and more libraries have adopted CNMARC for cataloging of Chinese materials (Chen, “Zhong wen tu shu ji du mu lu zhu ti biao yin cai yong hou kong gui fan de she xiang”).

In Taiwan Province, a local MARC format (later called CMARC) for monographs, which was strictly based on UNIMARC, was published early in 1981 and was revised almost at the same year, while

a MARC format for all types of publications was published in July 1984. This format was earlier than the CNMARC in the Chinese mainland, and was used as an example for drafting of the CNMARC (Chen, “Da lu he Taiwan de Zhong wen MARC bi jiao”). Although both are based on UNIMARC, CNMARC and CMARC are slightly different, especially in the use of character sets and terminologies (Zhu and Mengjie).

In January 2002, *China MARC Format / Authorities* was released as an industrial standard of the Ministry of Culture (WH/T15-2002). It was based on UNIMARC / Authorities published by IFLA in 1991. Since then, there has been no revision of the format (*China MARC Format / Authorities*).

## 4 UNIMARC Translations

UNIMARC Bibliographic Format was translated into Chinese in 1986. However, it was just used for internal use for the drafting of CNMARC, and was not published and distributed through commercial channel (Chen, “Zhongguo ji du mu lu tong xun ge shi cun zai de wen ti yu xiu gai jian yi”). There have been translations of some later editions, but none of them was published.

Translation is very important for correct understanding of the original UNIMARC text. In *Manual of the New Edition of CNMARC* in March 2004 (Guo jia tu shu guan, *Manual of the New Edition of CNMARC*), for example, EAN (073 - International Article Number) was misunderstood as a kind of numbers assigned to articles in published journals.

By the end of 2013, the Chinese translation of *UNIMARC Manual: Authorities Format* (3rd Edition, 2009) will be published by the National Library of China Publishing House. The translation was done by the Working Group for the Drafting of *CNMARC Authorities*

*Format*, a national standard proposal supervised by the National Library Standardization Technical Committee in 2009 and approved by the Standardization Administration of the People's Republic of China.

## 5 CNMARC Revisions

After the publication of CNMARC in 1991, there were some critics on it, pointing that:

- There are too many fields and subfields.
- There are some duplicated fields or subfields.
- The definitions of some data elements are not clear.

In March 1993, the Ministry of Culture approved the proposal of CNMARC Format as a library industrial standard (Qi). The standard (WH/T0503-96) was published in 1996, and began to be implemented in July 1, 1997 (Zhu). This standard was based on the 1994 edition of *UNIMARC Manual*. At the same time, main contents of the standard were included in the published *China MARC Format Manual* (1995) (“Zhongguo ji du mu lu ge shi ji shi yong shou ce nian nei chu ban”), and the revised edition of the manual was published in 2001 (Pan).

The industrial standard (WH/T0503-96) was revised according to the latest edition (2002) of UNIMARC, and was intended to become a national standard. The draft for the national standard was completed and approved by an expert committee organized by the Ministry of Culture in 2003, and its major contents were published as *Manual of the New Edition of CNMARC* in March 2004. Because of some procedural problems, the national standard was not finally released.

In December 2009, the Standardization Administration of the People's Republic of China approved the proposal of two national standards, *CNMARC Bibliographic Format* and *CNMARC Authorities Format*.<sup>1</sup> The two national standards were planned to be released in 2011, but they are delayed due to personnel and technical problems.

## 6 Special Characteristics in CNMARC

In addition to some minor changes to UNIMARC format, the major differences of CNMARC are reflected in the definition of some local fields 9-, -9-, -9, and subfields, \$9 (and \$A, \$B, etc.).

For example (Bibliographic format):

- 091 Union Books and Serials Number: numbers assigned by Chinese Administrative Agencies for books and serials, especially before the implementation of ISBN and ISSN in China.
- 092 Order Number: Numbers assigned by distributors.
- 094 Standard Publication Number: For the numbers of international, national, industrial or enterprise standards.
- 191 Coded Data Field: Rubbings
- 192 Coded Data Field: Ethnic Music of China
- 193 Coded Data Field: Chinese Antiquarian – General
- 194 Coded Data Field: Chinese Antiquarian – Copy Specific Attributes
- 393 Outsystem Chinese Character Note: For the description of Chinese characters not defined in character set.

---

<sup>1</sup>Standardization Administration of the People's Republic of China, December 14, 2009: [http://www.sac.gov.cn/gjbzjh/201012/t20101213\\_56788.htm](http://www.sac.gov.cn/gjbzjh/201012/t20101213_56788.htm).

- 690 Chinese Library Classification
- 692 Classification for Library of Chinese Academy of Sciences
- 696 Other Local Class Numbers
- 905 Holding Information: Used in the second edition of the *China MARC Format Manual* (2001), when Chinese libraries did not use separate MARC format for holding records.

As to Pinyin Romanization, NLC and public libraries use subfield \$9 to include all Latin scripts corresponding to Chinese characters in major subfields of the field, while CALIS (China Academic Library & Information System) and university libraries use \$A, \$B etc. for Latin scripts corresponding to Chinese characters in \$a, \$b etc.

Besides, CNMARC doesn't include some UNMARC fields not applicable to Chinese library cataloging, e.g. 012 (Fingerprint Identifier) and 670 (PRECIS).

As libraries in China do not use the concept of main entry for the cataloging of monographs in Chinese language, fields 700, 710, 720 are usually not used. Besides, name/title access points are not used in NLC records. Therefore, there are difficulties in adding uniform titles in bibliographic records, and there are very few records with uniform titles.

In NLC records, 701\$b and 701\$g are not used, while in CALIS records, 701\$b and 701\$g are used for foreign names (Xie).

Punctuations are not used in CNMARC records, and they are created automatically by computer systems in OPAC display. However, there are some problems with the punctuations. For example, 200\$d is for Parallel Title Proper and 200\$e is for Other Title Information, and we are hard to define automatic generation of punctuations for parallel other title information, because there are many possibilities for 200\$e. However, CALIS solved this problem by adding punctuations in these particular subfields.

## 7 Co-Existence of Two MARC Formats

In China, libraries use two MARC formats, MARC21 and CNMARC. CNMARC has some special fields for Chinese publications in addition to UNIMARC fields. Most small-sized libraries use UNIMARC-based CNMARC for all materials, with the consideration of easy management. For large-sized libraries that have sizeable collections in foreign languages, they prefer USMARC/MARC21 to CNMARC with the consideration of international compatibility and easy copy cataloging.

In the National Library of China, we use CNMARC, Chinese Library Classification and Classified Chinese Thesaurus for Chinese publications, and we use MARC21 for foreign publications, including those in Western languages, Japanese and Russian. For publications in Western languages, we use AACR2 (later RDA), LC Subject Headings, LC Name Authority File and Chinese Library Classification. The Aleph500 system allows us to maintain two separate databases respectively in CNMARC and MARC21 formats. There is no relationship between the two databases now. We were considering the possibility of establishing relationships between the two authority databases respectively in CNMARC and MARC21, but the cost was estimated very high.

In the Chinese mainland, most libraries use UNIMARC-based CNMARC for Chinese publications. In Hong Kong Special Administrative Region, most libraries follow the practices of the Western countries and use MARC21 for all publications. In Macau Special Administrative Region, some libraries use CNMARC and some use MARC21. Large libraries in Taiwan Province use CMARC (also based on UNIMARC and similar to CNMARC) or MARC21; mid-sized libraries favor CMARC. For libraries using UNIMARC-based formats, they also have different rules, especially for name headings and authority records.

As to Romanization, most libraries use Pinyin now, but they sometimes use different ways of segmentation, for example "mao ze dong" or "Mao Zedong" (Gu).

## 8 Holdings Format

At the early stage of library automation in China, there was no local library automation system using the technology of relational database, and there was no need to use a separate holdings format. Libraries used tag 905 in the bibliographic format to record holdings data.

Since the beginning of the 21th century, more and more libraries use international systems, such as Aleph 500, and people have to consider the use of holdings format. Because there was no UNIMARC holdings format, a research group of the National Library of China considered the localization of MARC21 Format for Holdings Data. In 2003, a Chinese translation of the MARC21 Format for Holdings Data was published (Guo jia tu shu guan, *MARC21 Format for Holdings Data*), and the format began to be applied in NLC and other libraries.

## 9 Problems in International Exchange

CNMARC is used mainly for the distribution of MARC records in China. Sometimes, we have information exchange with East Asian libraries, and CNMARC records are easy to be converted into JapanMARC (Japan) and CMARC (Taiwan Province) records, which are based on UNIMARC. However, we have difficulties in sharing bibliographic information with libraries in Hong Kong, a special administrative region of China, which use MARC21, and

those in Republic of Korea, which use KorMARC based on MARC21. Further, we have more difficulties in sharing CNMARC records with foreign libraries, most of which use MARC21 format.

After almost ten-year discussions and negotiations, NLC and OCLC signed an agreement for batch uploading of NLC Chinese bibliographic records to OCLC WorldCat database in 2008,<sup>2</sup> and had uploaded 2.3 million records by the end of 2008.<sup>3</sup> This is a great step forward for Chinese MARC records to be used by foreign libraries. However, because exact CNMARC-MARC21 automatic conversion is almost impossible, and manual conversion requires huge costs of human resources, NLC and OCLC had to choose the option of batch conversion, leaving details to be upgraded by OCLC users. The batch uploading process also adds NLC holding information to WorldCat.

## 10 Future Perspectives

Although there have been talks about the death of MARC, CNMARC is still a major format for the cataloging of Chinese materials and exchange of bibliographic records in China. If there is a need for records with other formats, such as XML, we just convert CNMARC into them. There is not any cataloging agency distributing XML records yet.

Since the bibliographic year 2009, the German National Library has been delivering its data in the format MARC 21.<sup>4</sup> Some libraries

---

<sup>2</sup> National Library of China to Add Its Records to OCLC WorldCat: <http://newsbreaks.infotoday.com/Digest/National-Library-of-China-to-Add-Its-Records-to-OCLC-WorldCat-41153.asp>.

<sup>3</sup>[http://www.nlc.gov.cn/dsb\\_zx/gtxw/201201/t20120116\\_58355.htm](http://www.nlc.gov.cn/dsb_zx/gtxw/201201/t20120116_58355.htm).

<sup>4</sup>MARC21, Deutschen Nationalbibliothek: [http://www.dnb.de/EN/Standardisierung/Formate/MARC21/marc21\\_node.html](http://www.dnb.de/EN/Standardisierung/Formate/MARC21/marc21_node.html).

in Taiwan Province also switched its bibliographic format from UNMARC-based CMARC to MARC21 in recent years.<sup>5</sup>

Chinese librarians have to consider the following questions:

- What is the original purpose for us to choose UNIMARC for the basis of CNMARC twenty years ago?
- Shall we still use UNIMARC for the basis of our cataloging?
- Shall we even abandon MARC formats and use a “modern” format?

The author has answered the first question in this article, but he cannot answer the other two questions now. The author doesn't think there will be any answers in the next few years. CNMARC will still be our major format for the cataloging of printed resources. Some new fields and subfields related to FRBR in the updates 2012 will be good tools for the application of FRBR in Chinese libraries.

---

<sup>5</sup><http://catwizard.net/posts/20120125095904.html>.

## References

- An, Shulan. "MARC ci dai yu wo guo xi wen bian mu". *Computers and Libraries* 4. (1983): 7–10. (Cit. on p. 265). Print.
- . "Yin jin MARC ci dai de shou ci ying yong". *Computers and Libraries* 4. (1981): 58. (Cit. on p. 265). Print.
- "Beijing tu shu guan xiang guo nei wai fa xing ji du mu lu". *Science and Technology Daily* February 21. (1991). (Cit. on p. 267). Print.
- Chen, Fuliang. "Da lu he Taiwan de Zhong wen MARC bi jiao". *Library Work and Research* 1. (1994): 28–30. (Cit. on p. 268). Print.
- Chen, Yuanzheng. "Beijing di qu xi wen tu shu ji du mu lu yan zhi jin zhan". *Library and Information Service* 5. (1983): 1–6, 41. (Cit. on p. 265). Print.
- . "Zhong wen tu shu ji du mu lu zhu ti biao yin cai yong hou kong gui fan de she xiang". *Journal of Library Science in China* 4. (1991): 69–71. (Cit. on p. 267). Print.
- . "Zhongguo ji du mu lu tong xun ge shi cun zai de wen ti yu xiu gai jian yi". *Library Development* 5. (1992): 53–57. (Cit. on p. 268). Print.
- China MARC Format / Authorities*. Beijing: Zhonghua ren min gong he guo wen hua bu fa bu, 2002. (Cit. on p. 268). Print.
- Gu, Ben. "National Bibliographies: the Chinese experience". *Alexandria* 18.3. (2006): 173–178. (Cit. on p. 273). Print.
- Guo jia tu shu guan. *Manual of the New Edition of CNMARC*. Pechino: Beijing tu shu guan chu ban she, 2004. (Cit. on p. 268). Print.
- . *MARC21 Format for Holdings Data*. Pechino: Ke xue ji shu wen xian chu ban she, 2003. (Cit. on p. 273). Print.
- IFLA. "Library of Congress Implements UNIMARC". *IFLA Journal* 11.3. (1985): 273–274. (Cit. on p. 266). Print.
- Liu, Rong. "Tan LC-MARC zai wo guo de shi yong". *Library and Information Service* 4. (1982): 22–23. (Cit. on p. 266). Print.
- Pan, Taiming. *China MARC Format Manual*. Pechino: Ke xue ji shu wen xian chu ban she, 2001. (Cit. on p. 269). Print.
- Qi, Siyan. "CN-MARC de wen shi ji she hui xiao ying". *Library Theory and Practice* 4. (1997): 29–30. (Cit. on p. 269). Print.
- Sun, Beixin. "The role of the National Library of China in standardization activities". *Scholarly Information and Standardization: Proceedings of the Twelfth Open Forum on the Study of the International Exchange of Japanese Information and Scholarly Databases in East Asian Scripts 1992/1993, November 20*. Tokio, 1992. (Cit. on p. 267). Print.
- Xie, Qinfang zhu bian. *Manual for CALIS Online Cooperative Cataloging, Part I*. Pechino: Beijing da xue chu ban she, 2000. (Cit. on p. 271). Print.

- “Zhongguo ji du mu lu ge shi ji shi yong shou ce nian nei chu ban”. *New Technology of Library and Information Service* 6. (1995): 64. (Cit. on p. 269). Print.
- “Zhongguo ji du mu lu ge shi xue shu yan tao hui ji yao”. *New Technology of Library and Information Service Suppl.* (1989): 10–55. (Cit. on p. 267). Print.
- Zhu, Qingqing and Huang Mengjie. *A Comparative Study of Chinese Cataloging Practices between Libraries across the Taiwan Straits*. Pechino: Shi jie tu shu chu ban gong si Beijing gong si, 2011. 15–47. (Cit. on p. 268). Print.
- Zhu, Yan deng qi cao. *CNARC Format*. WH/T0503-96 1996.2.6 fa bu 1997.7.1 shi shi. 1996. (Cit. on p. 269). Print.

BEN GU, Director of Foreign Acquisitions & Cataloging Department,  
National Library of China, Beijing, China.  
[bgu@nlc.gov.cn](mailto:bgu@nlc.gov.cn)

---

Gu, B. “MARC Formats in China: Local or International?”. *JLIS.it*. Vol. 5, n. 1 (Gennaio/January 2014): Art: #9083. DOI: [10.4403/jlis.it-9083](https://doi.org/10.4403/jlis.it-9083). Web.

ABSTRACT: The application of MARC formats in Chinese libraries has a history of about 20 years. At the age of internationalization and digitization, people ask the question: why we chose UNIMARC for the basis of CNMARC? Is it the best choice for China? Are there any other options? In this paper, the author reviews many historical documents and analyzes the present status of MARC formats in China.

KEYWORDS: Bibliographic formats; China; CNMARC; MARC; MARC21; UNIMARC.

---

Submitted: 2013-09-20  
Accepted: 2013-10-26  
Published: 2014-01-01







# JLIS.it



Dipartimento SAGAS, Storia, Archeologia, Geografia, Arte e Spettacolo

con il supporto di:

supported by:

**Casalini**  
**libri**



Le Lettere

La piattaforma ICT, lo sviluppo e la manutenzione dell'installazione di OJS che ospita JLIS.it sono forniti da:

ICT platform, developing and maintenance for the OJS installation hosting JLIS.it are provided by:



Direttore Responsabile ai termini di legge: Nicola Cavalli

In attesa di iscrizione nel registro stampa del Tribunale di Milano.

Finito di stampare nel mese di febbraio 2014 da

**Ledizioni**   
*The Innovative LEDpublishing Company*

<http://www.ledizioni.it>