

MetaLib e SFX: uno sguardo critico

Mario Ceroti

A proposito di metamotori di ricerca, link resolver e gestione delle risorse elettroniche negli atenei italiani

*Biblioteca Circolo giuridico
Università degli studi di Siena
ceroti@unisi.it**

Con questo articolo si intendono fornire alcune osservazioni relative all'usabilità e all'utilità di due strumenti sviluppati ed allestiti per la gestione della biblioteca digitale, adottati attualmente da diversi atenei italiani,¹ ed integrare quanto già evidenziato nei saggi scritti sull'argomento in cui vengono esaminati i pregi dei due software.² La nostra analisi, che si concluderà con una riflessione teorica, generale sulla gestione delle risorse digitali all'interno dei sistemi bibliotecari degli atenei italiani, non riguarderà la tecnologia sottesa ai due strumenti, bensì la loro capacità di soddisfare le esigenze informative, bibliografiche dell'utenza istituzionale in maniera amichevole ed esaustiva senza che questa debba ricorrere a personale specializzato, ai professionisti dell'informazione, nel rispetto di uno dei principi cardine della biblioteca digitale: la disintermediazione. MetaLib (prodotto da Ex Libris) viene proposto come strumento che permette di effettuare con un'unica interrogazione ricerche raffinate ed accurate all'interno delle risorse elettroniche disponibili in un ateneo, in sostanza come un metamotore di ricerca (*metasearcher*) da utilizzare nell'ambito delle collezioni digitali accademiche o, per citare le pagine web di presentazione e le guide in linea delle università

che lo hanno adottato, come uno strumento che "permette di eseguire una ricerca in diverse risorse simultaneamente" (Napoli Federico II); "effettuare una ricerca simultanea su più risorse, con un'unica interfaccia di interrogazione" (Padova); "effettuare una metaricerca (detta anche *integrated searching*, *federated searching* o *cross-database searching*), cioè di sottoporre la query a più risorse contemporaneamente" (Brescia); "ricercare simultaneamente, attraverso una semplice interfaccia di interrogazione, molteplici risorse sia acquistate dall'Università sia disponibili gratuitamente in rete e selezionate dai bibliotecari" (Siena); una "metaricerca in più banche dati" (Firenze).³

In realtà, la funzione di ricerca simultanea all'interno di risorse elettroniche eterogenee (aggregatori, banche dati, periodici elettronici, editori ecc.) è fortemente condizionata dalla limitazione numerica di risorse che possono essere contenute nei set disciplinari predefiniti nella maschera "Ricerca guidata" o che possono essere interrogate all'interno delle categorie o sottocategorie nella funzione "Scelta e ricerca".⁴ In breve: delle risorse elettroniche che un ateneo rende disponibili, solo una piccola parte è ricercabile mediante MetaLib sia per la succitata limitazione

numerica sia perché molte risorse non sono "metalibizzabili", ossia compatibili con le tecniche di ricerca di MetaLib. Occorre precisare che alcune guide in linea segnalano l'esistenza di questo limite numerico di risorse interrogabili: quella dell'Università di Firenze puntualizza che "la Ricerca guidata è un modo veloce e facile per ottenere risultati, ma non sempre è il modo migliore per eseguire la ricerca, visto che la ricerca tramite questa funzionalità è disponibile solo per un numero limitato di argomenti e di risorse" e che "purtroppo in alcune banche dati la ricerca è possibile solo visitando direttamente la risorsa. Queste banche dati, in particolare quelle italiane (fra cui poche sono ancora interrogabili via MetaLib), sono importanti quanto quelle nelle quali è possibile la metaricerca", mentre la guida in linea dell'Università di Siena avverte che "il numero massimo di risorse che MetaLib è in grado di interrogare simultaneamente è dieci" e che "molte risorse sono in lingua inglese". L'efficacia nel recupero delle informazioni di MetaLib, infatti, è condizionata anche dall'aspetto linguistico del contenuto dei metadati dei documenti, il che limita pesantemente – in assenza o quasi di descrizione semantica a monte – l'eshaustività delle ricerche compiute. In so-

stanza, MetaLib rischia di creare un'illusione di esaustività della ricerca bibliografica nell'utenza istituzionale (docenti e studenti) che, convinta di avere recuperato tutti i documenti utili relativi ad un determinato argomento, in realtà viene privata di informazioni bibliografiche contenute nelle più importanti risorse elettroniche disciplinari.

A conferma di ciò possiamo produrre prove ed esempi concreti relativi a tre fra i maggiori ambiti disciplinari accademici: l'italianistica, le discipline giuridiche e quelle di ambito biomedico.

Per quanto riguarda il primo settore disciplinare, MetaLib non fornisce una ricerca sufficientemente esaustiva poiché la più importante ed utilizzata risorsa bibliografica elettronica (LIRA, Letteratura italiana repertorio automatizzato) non è interrogabile tramite il metamatore, non è "metalibizzabile". Pure per l'altra importante risorsa bibliografica elettronica nell'ambito dell'italianistica (Italinemo), abbiamo un problema analogo: anche se interrogabile tramite l'interfaccia di MetaLib, i risultati della ricerca non vengono forniti contestualmente a quelli recuperati dalle altre risorse elettroniche, ma sono visualizzati nel sito della risorsa nativa a cui l'utente viene dirottato mediante un collegamento ipertestuale. In sostanza si tratta di un link alla risorsa nativa dissimulato o meglio dilazionato. Anche in questo caso MetaLib non adempie a pieno al suo compito: quello di interrogare simultaneamente le risorse elettroniche disponibili e fornire un'altrettanto simultanea visualizzazione del risultato della ricerca all'utente. L'indicizzazione di MetaLib non comprende neppure le riviste in formato elettronico disponibili attualmente in questa disciplina come "Testo", "Rivista di letteratura ita-

liana", "Studi novecenteschi", "Italianistica", "Moderna".

La situazione si ripropone in maniera analoga per le discipline giuridiche: una delle maggiori e più importanti banche dati di ambito giuridico, "De Jure", non è ricercabile tramite MetaLib, ma è interrogabile mediante l'interfaccia della risorsa nativa e così accade per "West Law International", "InfoLEGES", "Foro italiano online", "Il Sole24 ore" (banca dati). Per le discipline del settore biomedico, occorre dire che risorse elettroniche come TOXNET (ricchissimo portale che raccoglie basi dati bibliografiche e fattuali di chimica e tossicologia) non è ricercabile tramite MetaLib, mentre la principale banca dati di ambito medico, PubMedO, è interrogabile tramite il metamatore ma la ricerca eseguita all'interno della risorsa nativa è infinitamente più raffinata e veloce ed i risultati sono molto più "elaborabili", grazie anche alla presenza di numerosi filtri nell'interfaccia: in sostanza MetaLib si dimostra uno strumento poco efficace ed adeguato per l'interrogazione di PubMed che consente strategie di ricerca estremamente più raffinate se utilizzato mediante l'interfaccia della risorsa nativa. Altra risorsa elettronica disciplinare la cui "metalibizzazione" risulta alquanto problematica è la Cochrane Library (banca dati bibliografica e fattuale di secondo livello per la EBM, Medicina basata sull'evidenza, raccoglie revisioni sistematiche della letteratura, *trials* clinici ecc.) che presenta lo stesso inconveniente riscontrato per Italinemo: la ricerca è eseguibile tramite MetaLib ma i risultati vengono visualizzati all'interno della risorsa nativa.

Un'altra lacuna di MetaLib, relativa all'efficacia nel recupero dei documenti e che può generare eccessive aspettative nell'utenza

riguardo all'esaustività della ricerca bibliografica, è rappresentata dalla mancata indicazione su dove agisce la ricerca del metamatore: se all'interno del documento in formato elettronico, nel full text o sul contenuto dei metadati, dei record bibliografici predisposti per il recupero dell'informazione.

Per quanto riguarda poi l'usabilità e l'amichevolezza, abbiamo riscontrato alcune pecche nella navigabilità all'interno dei risultati della ricerca: difficoltosa, infatti, si dimostra l'operazione di recupero dell'elenco delle occorrenze della ricerca effettuata una volta visualizzato il record in formato completo. Utilizzando, infatti, il pulsante "Indietro" del browser l'utente non torna all'elenco delle occorrenze recuperate bensì al record, al metadato precedente: si ha in sostanza una visualizzazione sequenziale dei record. E così cliccando sulla funzione "Precedente" e "Successiva": lo scorrimento riguarda il record e non la pagina. L'unica soluzione per fare ritorno all'elenco dei risultati – in mancanza di un pulsante preposto a ciò con la dicitura che espliciti chiaramente questa funzione – è quella di selezionare uno dei due formati di visualizzazione dei metadati ("Formato tabulare" o "Formato citazione"). Anche la navigazione tra registrazioni, tra metadati risulta pressoché inesistente, non essendo disponibili collegamenti ipertestuali tra record né partendo dall'intestazione né dai descrittori o notazioni di una qualsiasi classificazione bibliografica. Per quanto riguarda poi la restituzione dei risultati di una ricerca, il programma opera un filtraggio automatico, effettua una clusterizzazione, una faccettazione delle occorrenze, ma non permette all'utente di predisporre, di definire filtri particolari né

di affinare le ricerche compiute. L'altro strumento, spesso abbinato a MetaLib, è SFX (software sviluppato per la gestione dei collegamenti citazionali o *link resolver*)⁵ basato sulla tecnologia OpenUrl, *context sensitive* e che dovrebbe permettere all'utente la navigazione ipertestuale tra documenti elettronici full text, nonché un'integrazione fra le risorse digitali e il catalogo elettronico. In realtà la navigazione tra documenti non è praticabile: non è possibile accedere al testo pieno di un articolo citato in nota all'interno di un altro documento, in quanto non esistono collegamenti ipertestuali tra citazioni bibliografiche. Anche l'integrazione con il catalogo bibliografico risulta priva di una sua utilità, se limitata ad un numero esiguo e marginale di documenti: il catalogo elettronico solitamente non contiene, se non in minima parte, record relativi ai titoli analitici a testo pieno e quindi il collegamento citazionale, l'indirizzamento al full text, non avviene mediante la ricerca effettuata sull'OPAC. L'accesso al testo pieno, inoltre, non è così agevole neppure da una registrazione recuperata tramite MetaLib, in quanto spesso SFX non offre immediatamente il full text, ma dirotta l'utente sulla pagina della risorsa nativa (aggregatore, editore, e-journal, banca dati) in cui ricercare il microdocumento a testo pieno. Il metadato del microdocumento inoltre contiene, oltre al pulsante SFX, un link denominato "Link esterno", che il più delle volte indirizza l'utente all'interfaccia della risorsa nativa e non al testo pieno dell'articolo. SFX permette anche l'accesso ad alcuni servizi bibliografici come la ricerca automatica in cataloghi elettronici esterni (Mai Azalai, ACNP, SBN) o la formattazione di un record bibliografico nello stile citaziona-

le prescelto. Spesso accade però che la citazione bibliografica ottenuta sia incompleta, lacunosa e non corrispondente allo stile citazionale indicato per quanto riguarda l'ordine, la sequenza degli elementi bibliografici. Anche la ricerca automatica su OPAC esterni spesso non ha un esito positivo, poiché il servizio di SFX frequentemente dirotta all'interfaccia del catalogo da interrogare senza però riprodurre la stringa di ricerca in maniera automatica. Da quanto evidenziato possiamo dedurre che i due strumenti, pur dimostrandosi di una certa utilità per l'utente di una biblioteca digitale, non adempiono a pieno alla loro funzione, poiché MetaLib in quanto metamatore si dimostra fortemente limitato dal numero di risorse elettroniche interrogabili simultaneamente e dall'esiguità di macrodocumenti digitali metaricercabili, mentre SFX non permette il più delle volte un accesso diretto al full text del documento o la navigazione ipertestuale tra citazioni bibliografiche e testo pieno, né una vera e propria integrazione fra catalogo elettronico bibliografico e risorse digitali.

Per concludere, vorremmo fare una considerazione generale ed offrire una veloce panoramica sulle modalità di gestione delle risorse elettroniche nel contesto accademico italiano. La maggior parte degli atenei ricorre all'utilizzo di portali con elenchi *A to Z* e cataloghi separati per le risorse elettroniche; altri hanno adottato software per l'integrazione bibliografica o *link resolver*, ma nessuno propone l'accesso alle collezioni digitali tramite quelli che Mauro Guerrini e Riccardo Ridi hanno definito rispettivamente "accesso principale alle risorse remote"⁶ e il "punto di accesso fondamentale a una biblio-

teca digitale”,⁷ lo “strumento principe” per l’integrazione bibliografica tra catalogo elettronico e risorse elettroniche, ovvero l’OPAC. In sostanza, dell’originale soluzione proposta da Riccardo Ridi riguardo alla gestione delle collezioni digitali il contesto accademico italiano preferisce privilegiare la logica della preservazione rispetto a quella dell’accesso ed adottare una soluzione di basso profilo: mantenere l’OPAC tradizionale (o *small OPAC*) ristretto ai soli documenti posseduti ed utilizzare per l’accesso alle RER liste di archivi separati dall’OPAC (o VRD). Gli atenei italiani propendono, nella gestione delle collezioni digitali, per la logica della preservazione, evitando soluzioni come l’allestimento di un OPAC allargato o portalizzato o esteso (o *big OPAC*) che “include tutti i documenti posseduti o selezionati dalla biblioteca”,⁸ pur nella consapevolezza che la sua realizzazione pratica comporterebbe un notevole impegno in termini di risorse umane ed economiche.⁹ Spesso, anzi, la fruizione delle risorse digitali viene resa particolarmente macchinosa dalla tendenza da parte degli atenei a mantenere separati i cataloghi o le semplici liste alfabetiche in base alla tipologia di risorsa (banche dati o periodici elettronici), frammentando di fatto la collezione digitale accademica e costringendo l’utente a ripetere le ricerche in differenti cataloghi. In sostanza, viene riprodotta nell’ambito delle risorse elettroniche la situazione relativa agli OPAC accademici che abbiamo analizzato e descritto in un nostro precedente articolo: la quarta legge di Ranganathan viene nuovamente disattesa o meglio, in questo caso, la sua “riformulazione sub specie ipertestuale”¹⁰ compiuta da Ridi nella definizione delle cinque

leggi della biblioteca ipertestuale – crea i link più diretti – nonché la purtroppo assai poco applicata “three click rule”.¹¹

Vorremmo cogliere l’occasione, inoltre, per compiere una sorta di simulazione, una dimostrazione virtuale di come dovrebbe essere, di come immaginiamo una biblioteca digitale compiutamente realizzata, concretizzata utilizzando il risultato bibliografico di un progetto – purtroppo poco conosciuto – attivato e concluso presso il Dipartimento di filologia e critica della letteratura dell’Università degli studi di Siena. Il progetto SELECT (Spoglio elettronico di letteratura italiana, comparatistica e teoria letteraria) ha portato alla creazione di una base di dati che comprende circa 15.000 record bibliografici relativi ad articoli apparsi su oltre quaranta riviste scientifiche negli ultimi quarant’anni (1970-1999) e che è confluita nel catalogo elettronico del Servizio bibliotecario senese. Dotando i metadati relativi ai titoli analitici, ai microdocumenti di un link al full text potremmo realizzare quanto teorizzato da Ridi e Guerrini – l’utilizzo dell’OPAC come punto di accesso principale alle risorse analogiche ed elettroniche di una biblioteca ibrida – riducendo a tre semplici operazioni il percorso ipertestuale che l’utente deve compiere per fruire della risorsa bibliografica (figure 1 e 2).¹² Riteniamo che la strada da percorrere per la realizzazione di un’effettiva ed efficace integrazione bibliografica tra catalogo elettronico e risorse digitali sia l’accentuazione dell’interoperabilità tra differenti archivi. Gli editori, i fornitori di servizi commerciali, gli sviluppatori di programmi per l’integrazione bibliografica dovrebbero garantire la possibilità di utilizzare i metadati – creati magari attenendosi agli standard

catalografici ed ai formati internazionali per l’interscambio dei dati bibliografici – consentendo la loro estrazione e derivazione all’interno dei cataloghi elettronici così da permettere una reale integrazione bibliografica. L’OPAC a questo punto rappresenterebbe lo strumento di ricerca principale per il recupero delle risorse analogiche e digitali grazie ai metadati in esso contenuti e comprensivi del localizzatore, dell’identificatore che indirizza l’utente al documento a testo pieno conservato in un archivio a sé stante: in sostanza, il catalogo elettronico andrebbe a costituire l’archivio che contiene le registrazioni catalografiche, i metadati, lo strumento di ricerca bibliografica principale, mentre un deposito a parte dovrebbe contenere i microdocumenti in formato elettronico full text a cui accedere tramite un collegamento ipertestuale inserito all’interno del metadato (una riproposizione *mutatis mutandis* dell’architettura, dello standard OAI con la distinzione fra *service provider* e *data provider*). Riteniamo infatti che sia in questa direzione – favorendo ed incrementando l’interoperabilità, la produzione di metadati rigorosi e di qualità e l’attivazione di procedure automatiche – che può essere vinta la sfida per la realizzazione di un’efficace e proficua integrazione bibliografica tra risorse digitali e catalogo elettronico.

Note

* Si ringrazia per la consulenza tecnica Antonella Carli (responsabile dell’Ufficio periodici della Biblioteca Circolo giuridico dell’Università degli studi di Siena). Le prove che abbiamo compiuto per la stesura di questo articolo sono state eseguite sui programmi MetaLib e SFX disponibili presso

l'Università degli studi di Siena e, dove possibile, anche sui software presenti negli altri atenei italiani.

¹ SFX (Università della Calabria, Salerno, Parma, Bologna, Cassino, Genova, Milano Bicocca, Milano Statale, Università dell'Insubria, Università della Toscana); MetaLib più SFX (Brescia, Firenze, Napoli Federico II, Padova, Sassari, Siena, Venezia Ca' Foscari).

² CINZIA BUCCHIONI, *SFX e OpenURL. Gli esperimenti del team di Van de Sompel*, "Bibliotime", 5 (2002), 2, <<http://didattica.spbo.unibo.it/bibliotime/num-v-2/bucchion.htm>>; EAD., *L'evoluzione dello standard OpenURL*, "Bibliotime", 7 (2004), 3, <<http://didattica.spbo.unibo.it/bibliotime/num-vii-3/bucchion.htm>>; EUGENIO PELIZZARI, *Il context-sensitive linking alla prova degli utenti. Uno studio sull'utilizzo di Siro-SFX presso l'Università degli studi di Brescia*, "Biblioteche oggi", 24 (2006), 6, p. 21-27; CINZIA BUCCHIONI – SERAFINA SPINELLI, *Teoria e pratica del linking citazionale*, "Biblioteche oggi", 25 (2007), 8, p. 70-86, <<http://www.bibliotecheoggi.it/2007/2007080701.pdf>>.

³ Firenze, <http://www.sba.unifi.it/upload/dl/Materiali_per_studenti/Metalib_guida_rapida.pdf>; Brescia <<http://www.unibs.it/on-line/bib/Home/ServiziAteneo.html>>; Siena, <<http://www.asb.unisi.it/asb/ita/intro.html>>; Padova, <http://aire.cab.unipd.it:8331/V/PYLM45VFX1ECSHLTHPPCJT2S6EKYSN9R3T4H288N3CDN178P6Q10310?&pds_handle=GUEST>; Napoli Federico II, <http://143.225.172.200:8331/V/HVJ5FX7R4QK2QR9SEV3LRPT6TRSRA24LVBA2S2LVJS3B1S74P-91433?func=help&file_name=META§ion=search>.

⁴ Il numero massimo di risorse interrogabili simultaneamente da MetaLib è dieci. Questo è il numero indicato nella Guida in linea consultabile presso l'Università di Siena. Le differenti strategie di ricerca disponibili in MetaLib vengono indicate dai singoli atenei con diciture differenti: ricerca veloce, metaricerca ecc.

⁵ HERBERT VAN DE SOMPEL – PATRICK HOCHSTENBACH, *Reference linking in un ambiente di biblioteca ibrida. Parte 1: Strutture per il linking*, trad. di Cinzia Bucchioni, <<http://www.aib.it/aib/commiss/cnur/trsomp1.htm#3>>; Id., *Reference linking in un ambiente*

di biblioteca ibrida. Parte 2: SFX, una soluzione generica di linking, trad. di Cinzia Bucchioni, <<http://www.aib.it/aib/commiss/cnur/trsomp2.htm#3>>; Id., *Reference linking in un ambiente di biblioteca ibrida. Parte 3: Generalizzazione della soluzione SFX nell'esperienza 'SFX@ Ghent & SFX@LANL'*, trad. di Cinzia Bucchioni, <<http://www.aib.it/aib/commiss/cnur/trsomp3.htm#3>>

Fig. 1

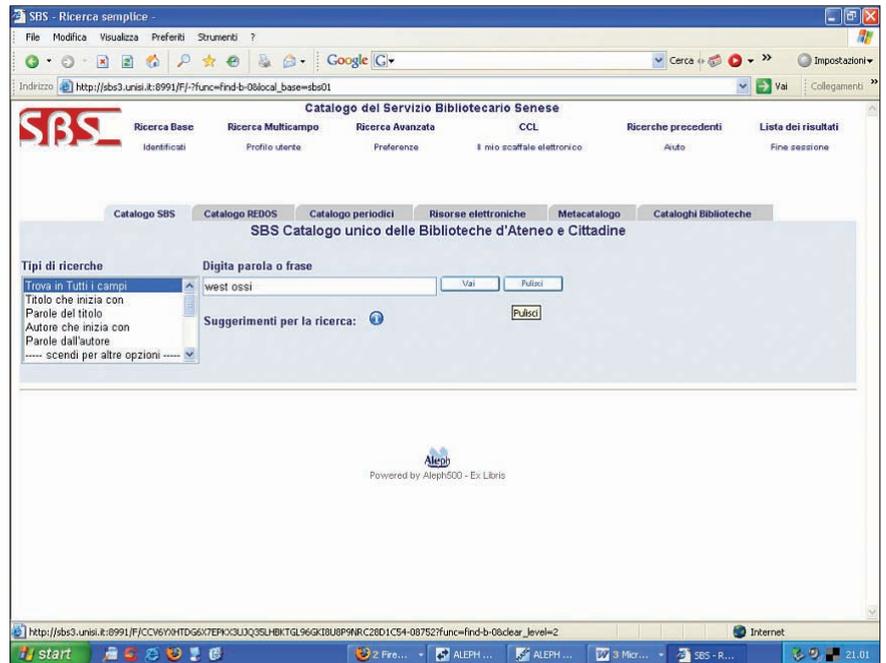
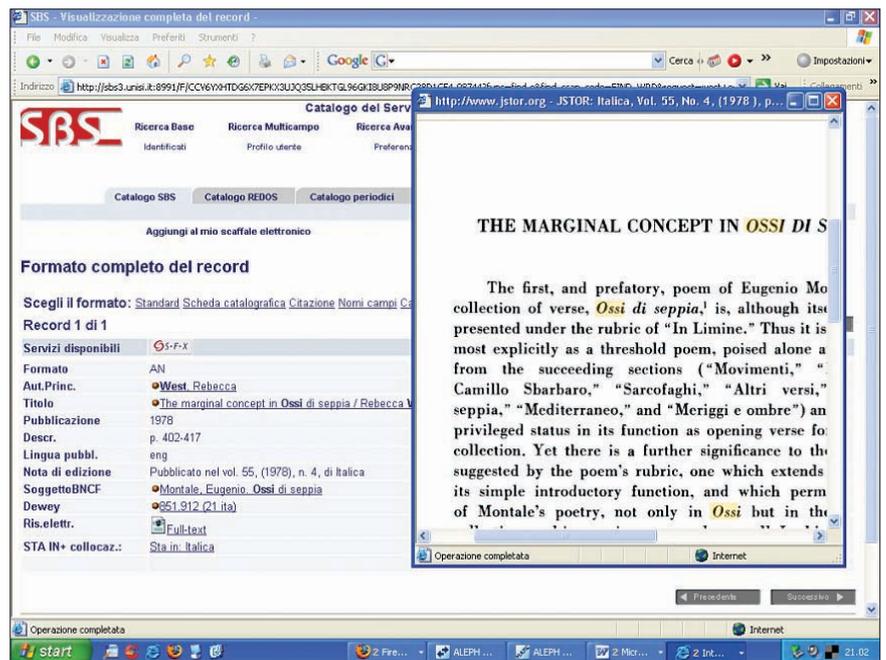


Fig. 2



>; CINZIA BUCCHIONI, *SFX e OpenURL: gli esperimenti del team di Van de Sompel*, "Bibliotime", 5 (2002), 2.

⁶ MAURO GUERRINI, *Il catalogo della biblioteca ibrida. Una rivoluzione copernicana*, "Biblioteche oggi", 20, (2002), 6, p. 44; Si veda anche sull'argomento il contributo di STEFANIA MANZI – ENRICO MARTELLINI, *Il catalogo e le risorse elettroniche. Un'integrazione possibile*,

“Bollettino AIB”, 43 (2003), 1, p. 7-28, <<http://www.aib.it/aib/boll/2003/03-1-007.htm>>.

⁷ RICCARDO RIDI, *La biblioteca come ipertesto*, Milano, Editrice Bibliografica, 2007, p. 117 e 201.

⁸ *Ibidem*, p. 119.

⁹ Per una panoramica sulla gestione delle risorse elettroniche in ambito accademico italiano si veda MARIA CASSELLA,

Le risorse elettroniche nei portali di biblioteche, “Biblioteche oggi”, 25 (2007), 10, p. 5-15, <<http://www.bibliotecheoggi.it/content/20071000501.pdf>>.

¹⁰ MICHELE SANTORO, *Biblioteca. Metafore e modelli*, “Bibliotime”, 10 (2007), 3; MARIO CEROTTI, *Ranganathan disatteso. Sull'usabilità e l'amichevolezza degli OPAC accademici italiani*, “Biblioteche oggi”, 25 (2007), 10, p. 16-21.

¹¹ JEFFREY ZELDMAN, *Taking your talent to the Web*, Indianapolis, New Riders, 2001.

¹² Abbiamo constatato che il pulsante che attiva le funzioni SFX non permette di accedere al testo pieno dell'articolo contenuto in JSTOR e quindi disponibile per l'utente che, partendo dal metadato relativo al microdocumento, riceve un'informazione scorretta e non attendibile riguardo alla possibilità di fruire del full text. Anche recuperando il record catalografico del macrodocumento – in questo caso la rivista “Italice” – e compiendo diversi passaggi ipertestuali, l'utente non riesce ad accedere al testo dell'articolo in questione. L'accesso al full text del documento non avviene direttamente tramite il link, l'identificatore contenuto nel suo metadato (anzi, attivando il pulsante SFX dalla registrazione catalografica relativa all'articolo abbiamo un'informazione fuorviante: che non esiste, non è disponibile il testo pieno del saggio), bensì dal metadato della risorsa elettronica che lo contiene, del documento ospite, del macrodocumento (figure 3 e 4) mediante un collegamento all'aggregatore – JSTOR – che contiene fisicamente il full text.

Fig. 3

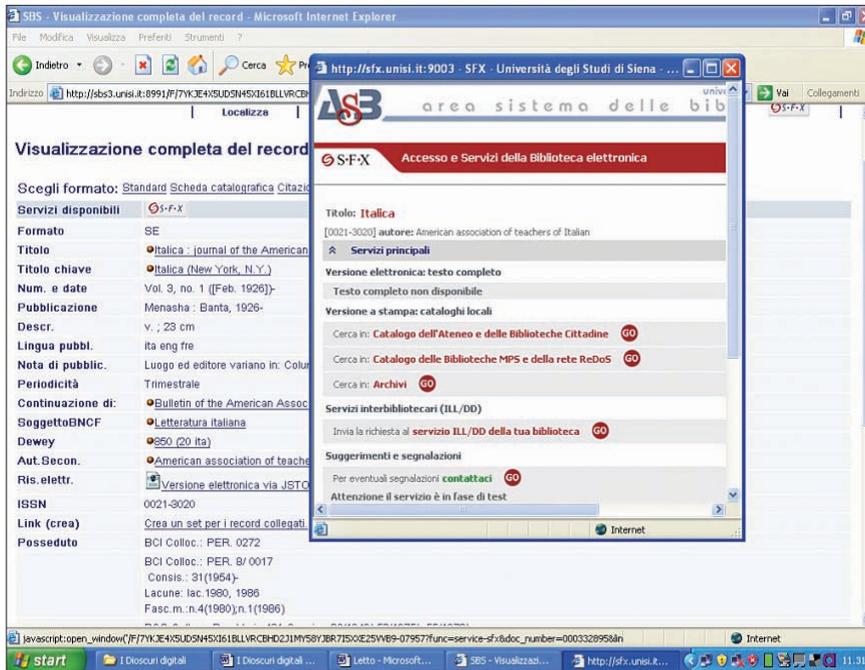
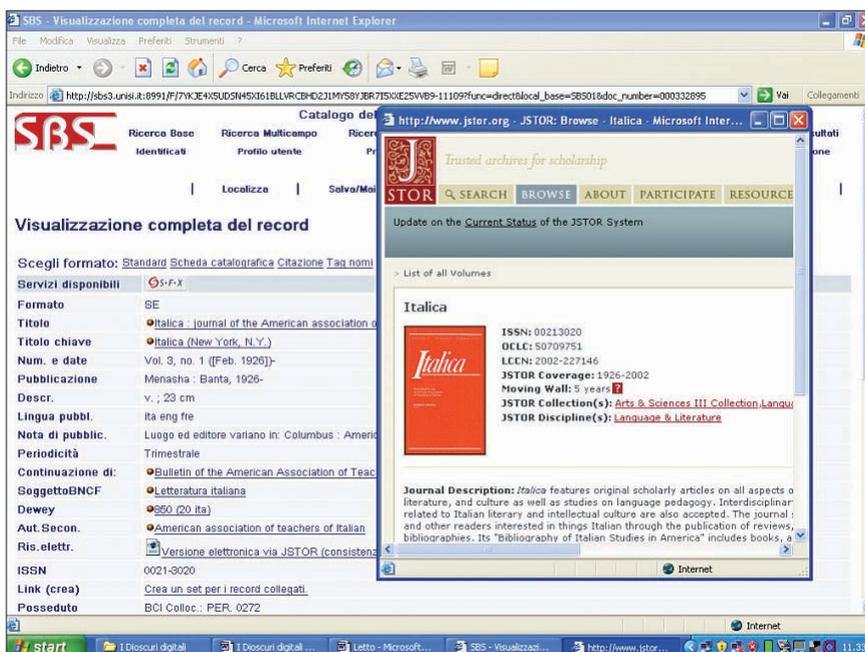


Fig. 4



Abstract

The essay includes an a parte utentis analysis of one of the most widespread tools of bibliographic integration – Metalib-SFX- and a short consideration on managing electronic sources in Italian universities, also considering Riccardo Ridi and Mauro Guerri- ni's theoretical contribution about the need to apply OPAC as the only access point to library's analogic and electronic sources. Moreover, the essay proposes a solution for a satisfying bibliographic integration that implies, by software companies, increasing implementation based on rigorous and high quality meta- data and automatic procedures for importing data from different archives.