Teoria e pratica del linking citazionale

Cinzia Bucchioni*
Università degli studi di Pisa
bucchioni@angl.unipi.it
Serafina Spinelli*
Università degli studi di Bologna
serafina.spinelli@unibo.it

1. Introduzione al linking citazionale

1.1. Il contesto

Che le biblioteche digitali siano organismi complessi, ambienti tendenti a un alto tasso di entropia, terreni dai confini labili e in perenne movimento, e in definitiva oggetti delicati da maneggiare per i bibliotecari e difficili da padroneggiare per gli utilizzatori, non è certo la *Vulgata* della biblioteconomia delle magnifiche sorti e progressive, ma è una percezione che trova ormai numerosi riscontri, non solo nella nostra prassi di lavoro quotidiana ma anche in una visione teorica sempre meno monolitica ("non una grande narrazione sistematica…") e sempre più dialettica ("… ma tante conversazioni tenute insieme da un linguaggio comune"[1]). Basta soffermarsi sulla maggior parte delle biblioteche digitali allestite dalle nostre istituzioni (l'esperienza cui ci rifacciamo è universitaria, ma crediamo che il discorso sia per molti aspetti estensibile alle biblioteche digitali di sistemi bibliotecari di altra natura istituzionale), per poter annotare caratteristiche come le seguenti:
– s*ono collezioni digitali, ma ognuna è integrata da una o più collezioni cartacee*: non si tratta solo di ricordare il familiare concetto di biblioteca ibrida, ma piuttosto di indagare le implicazioni (concettuali, tecniche, logistiche, di servizio ecc.) del rapporto con una, due, tre o *n* collezioni cartacee, probabilmente decentrate su un territorio più o meno vasto, ognuna con la sua storia, le sue caratteristiche, la sua identità, i suoi *stakeholders* ecc.;
– *per converso, le biblioteche digitali hanno tipicamente un "profilo lasco", un'identità debole*: benché sia universalmente affermato che la formazione e la gestione di una collezione digitale dovrebbero seguire un piano di sviluppo scientificamente garantito, e che come tutte le biblioteche anche quelle digitali dovrebbero rispondere alle esigenze di istruzione, studio, ricerca ecc. delle istituzioni che le sviluppano, sappiamo che in realtà sono fortemente condizionate da fattori come le politiche degli editori, l'attività dei consorzi di acquisto, le capacità economiche e contrattuali degli enti gestori ecc.;
– *sono tipicamente costituite da una grande varietà di tipologie documentarie*, diversi formati, protocolli, standard, interfacce, piattaforme, condizioni e modalità di accesso, organizzazione, indicizzazione, granularità ecc.;
– *presentano un certo numero di problemi inediti o di natura diversa da quelli delle collezioni cartacee*, rispetto ai quali non sempre i bibliotecari sono dotati di solidi ferri del mestiere: citiamo ad esempio i problemi di negoziazione delle licenze, di organizzazione e gestione degli accessi delle categorie di utenti, dei profili e delle autorizzazioni collegate, di gestione dei servizi che agiscono nella sfera dei diritti di proprietà intellettuale, di armonizzazione dei costi fra i diversi centri di spesa e della politica degli acquisti fra i vari centri decisionali, di trattamento, promozione e formazione all'uso, di monitoraggio e valutazione quantitativa e qualitativa, di controllo e manutenzione delle localizzazioni e delle consistenze, di conservazione e garanzia di accesso nel tempo ecc.;
– *le biblioteche digitali abbattono le barriere fisiche* (è il loro mestiere: tipicamente, *n* utenti possono accedere contestualmente allo stesso documento, in qualsiasi momento della giornata, da qualsiasi punto della rete, senza muoversi dalla propria scrivania e senza neanche sapere su quale server il documento si colloca fisicamente), ma non sempre sono in grado di evitare che si creino altre barriere, non solo di tipo tecnologico o contrattuale, ma anche determinate dalla mancanza di competenze da parte degli utenti finali, o semplicemente dalla maggior difficoltà di orientarsi, di non perdere il filo, nel mondo non sequenziale della rete;
– *sono fruibili di norma in rete, e in un contesto ipertestuale*;
– *sono "meno intermediate" rispetto alle collezioni cartacee*: anche se teoricamente tutti i documenti della collezione digitale dovrebbero essere organizzati, catalogati e indicizzati, e resi facilmente accessibili ai vari livelli e con le metodologie più appropriate, non sempre ciò accade, per vari motivi, fra cui l'esistenza di metodi di accesso alternativi ai cataloghi (tipicamente, le piattaforme commerciali degli editori, degli aggregatori ecc.) e la "temporaneità" del possesso (di norma, più che possesso, accesso su licenza) per larga parte della collezione, che determina la

* Benché la ricerca da cui questo articolo deriva sia stata condotta in stretta collaborazione e il testo nel suo complesso sia interamente condiviso dalle autrici, la parte 1 (*Introduzione al linking citazionale*) è stata materialmente scritta da Serafina Spinelli; la parte 2 (*L'OpenURL* resolver *dell'Università di Pisa*) da Cinzia Bucchioni; la parte 3 (*Dall'applicazione alla visione: prospettive e sviluppi*) e i box sono stati scritti a quattro mani. Ringraziamo gli amici e colleghi che hanno collaborato in vario modo alla ricerca, e che hanno letto le bozze di questo testo.

[1] AIB. Gruppo di studio sulle biblioteche digitali, *Manifesto per le biblioteche digitali*, 2005, <http://www.aib.it/aib/cg/gbdigd05a.htm3>.

natura instabile e "non incrementale" dei cataloghi delle risorse digitali. C'è chi dice che non sono meno intermediate, ma semplicemente "diversamente intermediate": anche ammettendolo, lo sono certamente in maniera meno organica, meno "catalografica", ma mediante strumenti vari e diversi che, nativamente, non colloquiano fra loro.
Queste ultime caratteristiche – la fruizione in un ambiente ipertestuale e la bassa intermediazione catalografica – hanno favorito recentemente lo sviluppo e l'affermazione di nuove categorie di strumenti per le biblioteche digitali, fra cui:
– i sistemi di ERM (Electronic Resources Management),[2] come strumenti in grado di gestire l'intero ciclo di vita delle risorse digitali, dalla selezione alla valutazione, i relativi flussi di lavoro e le interazioni con gli altri sistemi di gestione della biblioteca e dei suoi servizi;
– i portali, come piattaforme per organizzare l'accesso ai servizi e ai contenuti, anche in maniera personalizzata;
– gli strumenti di ricerca federata, specializzati nella ricerca simultanea su OPAC, repertori, banche dati e altri contenitori di informazioni bibliografiche e documentali (aggregabili anche in collezioni predefinite e/o personalizzabili), nella presentazione di risultati "trattati" (aggregati, deduplicati, ordinati…) e nell'erogazione di servizi al lettore (ad esempio, il salvataggio di ricerche e bibliografie);
– e infine, i servizi di *reference linking*, o linking citazionale, come strumenti capaci, a partire dalle informazioni contenute in una citazione bibliografica su un qualunque servizio web, di arrivare all'oggetto informativo cui essa si riferisce e ad eventuali ulteriori servizi correlati, creando un tessuto di integrazione delle risorse digitali dell'istituzione.

1.2. "Protostoria" del linking esteso

È proprio sul servizio di linking citazionale, realizzato mediante soluzioni tecnicamente innovative ma poggiato su solide radici concettuali, che intendiamo focalizzare il nostro discorso: a chi cerchi giustificazioni teoriche o "storiche" per saggiarne la rilevanza basti pensare che si tratta in fondo della convergenza fra una prassi di studio e ricerca vecchia di secoli, anzi per certi versi di millenni (l'uso della citazione da parte degli autori per riconoscere crediti e influenze, per ricondursi a scuole di pensiero e autorità; da parte dei lettori per allargare la propria conoscenza, per risalire alle fonti, per espandere la ricerca; più recentemente – ma ormai da diversi decenni – l'uso delle citazioni anche da parte dei bibliotecari per l'indicizzazione associativa, l'analisi citazionale, le valutazioni bibliometriche ecc.) e della più evidente caratteristica del mondo del web, cioè quella di creare associazioni e consentirne la navigazione mediante i legami ipertestuali.
E infatti, già a partire dalla fine degli anni Novanta,[3] una volta definito il problema nei suoi termini generali – integrare informazione secondaria e informazione primaria, ovvero arrivare "con un clic" da una citazione bibliografica all'oggetto informativo cui si riferisce[4] –, il dibattito si è piuttosto focalizzato sulla discussione di una serie di requisiti e opzioni di carattere sia tecnico che strategico, e anche su una possibile interpretazione del *reference linking* in chiave più estesa rispetto al semplice puntamento da una citazione al suo oggetto, o meglio alla copia dell'oggetto più appropriata per l'utente in questione. Possiamo sintetizzare le opzioni implementative del linking citazionale in tre coppie oppositive, sulle quali torneremo anche nel prosieguo del discorso:
– *linking interno o esterno*, cioè linking realizzato collegando reciprocamente solo i documenti interni a una singola collezione o piattaforma commerciale (interno), oppure puntando verso qualsiasi documento citato (esterno);
– *linking chiuso o aperto*, cioè linking che punta a una certa istanza (o anche a più istanze, ma comunque definite dalla risorsa di partenza, cioè quella che contiene la citazione) del documento citato (chiuso), oppure che, prevedendo che i link vengano forniti da soggetti diversi da quello che fornisce la citazione, consente di realizzare l'indirizzamento all'istanza cui l'utente ha diritto di accesso (aperto);
– *linking statico o dinamico*, cioè linking realizzato associando a priori, una volta per tutte, il link alla citazione, sia pur in molti casi calcolandone l'indirizzo con procedure più o meno automatizzate (statico), oppure computato "al volo" solo al momento in cui serve per passare dalla citazione al suo oggetto (dinamico).
Storicamente, le prime implementazioni del linking citazionale, realizzate da parte di fornitori o aggregatori di risorse digitali sin dalla metà degli anni Novanta, sono state di tipo "statico" (cioè realizzate mediante l'inclusione di link precalcolati e fisicamente "scritti" entro la citazione) e "interne" (cioè realizzate all'interno della collezione del fornitore, e quindi in grado di fornire legami solo fra risorse del medesimo fornitore o aggregatore, o fra fornitori o aggregatori che stipulassero accordi preventivi per il linking reciproco).[5] Alle difficoltà tecniche e organizzative della gestione di una collezione virtualmente unica e in-

[2] Per un'introduzione ai sistemi di ERM si veda l'articolo di Maria Cassella – Paola Gargiulo, *Il workflow delle risorse elettroniche remote*, "Biblioteche oggi", 24 (2006), 6, p. 46-58, <http://www.bibliotecheoggi.it/2006/20060604601.pdf>.
[3] Cfr. ad es. Steve Hitchcock, Les Carr, Wendy Hall, Stephen Harris, S. Probets, David Evans, David Brailsford, *Linking electronic journals. Lessons from the open journal project*, "D-Lib Magazine", 4 (1998), 12, <http://www.dlib.org/dlib/december98/12hitchcock.html>; Priscilla Caplan – William Y. Arms, *Reference linking for journal articles*, "D-Lib Magazine", 5 (1999), 7-8, <http://www.dlib.org/dlib/ july99/caplan/07caplan.html>; Herbert Van de Sompel – Patrick Hochstenbach, *Reference linking in a hybrid library environment*, "D-Lib Magazine", 5 (1999), 4 e 10.
[4] "The generic statement of the reference linking problem is, 'Given the information in a standard citation, how does one get to the thing to which the citation refers?'" (Priscilla Caplan – William Y. Arms, *Reference linking for journal articles*, cit.).
[5] Hitchcock riporta come esempi di servizi di linking interno, nel suo citato articolo del 1998, quelli delle collezioni di BioMedNet, HighwirePress e Institute of Physics. Si cominciano ad affermare più tardi, verso la fine degli anni Novanta, i prodotti finalizzati a gestire il *reference linking* verso il maggior numero possibile di risorse, come SilverLinker, un vero e proprio database di link a full text di articoli distribuiti da editori e aggregatori convenzionati, integrabile nelle banche dati SilverPlatter.

terrelata tramite le citazioni, si aggiungeva infatti la resistenza degli editori a consentire il raggiungimento diretto degli articoli, percepito come una pericolosa perdita di controllo sui contenuti: anche le realizzazioni di *reference linking* reciproco fra collezioni di fornitori diversi (si potrebbe in questo caso parlare di modello "ibrido", a metà strada fra "interno" ed "esterno") non si fondavano su sintassi pubbliche o standardizzate di *deep linking*, bensì su precisi accordi commerciali.[6]

Un'ampia riflessione condotta nel 1999 all'interno della comunità professionale[7] promuove il superamento dei vecchi modelli di *reference linking* "interni" e "chiusi" tipici dei fornitori commerciali, sostenendo modelli basati sullo sviluppo di identificatori dotati di tre caratteristiche, cioè la persistenza (stabilità), l'unicità (anche non assoluta, purché garantita all'interno del proprio "dominio")[8] e la capacità di risoluzione (ovvero la capacità di supportare la risoluzione dell'identificatore verso tutte le copie dell'opera che rappresenta, o verso il sottoinsieme delle copie appropriate per l'utente).

L'articolo di Caplan e Arms, che possiamo considerare la *summa* di questa riflessione, muove da una riformulazione in chiave più ristretta del problema del *reference linking*, una riformulazione che prende pragmaticamente atto del fatto che, quanto meno in un contesto di tipo scientifico-accademico, il focus è centrato sulla citazione degli articoli: "Data l'informazione contenuta in una citazione ad un articolo di periodico, come può l'utente arrivare dalla citazione ad una copia appropriata dell'articolo?".[9] Caplan e Arms rilevano che la soluzione del linking realizzato mediante il puntamento diretto alla URL di una copia dell'articolo citato, oltre a soffrire della tipica instabilità delle localizzazioni del webspace, non tiene conto di quelle che – con linguaggio FRBR – si possono chiamare entità bibliografiche di riferimento del reference citazionale. Se è vero infatti che di norma è l'opera l'entità di riferimento della citazione (anche se in alcuni casi l'autore può volersi riferire a una sua particolare manifestazione, ad es. la versione online e non quella cartacea, la versione html e non quella pdf, o viceversa), e la copia (localizzata) è l'entità di riferimento della destinazione del link, il meccanismo di linking deve essere fondato su un identificatore in grado di rappresentare l'opera, di individuarne tutte le possibili copie ("risoluzione multipla") e le relative localizzazioni. Solo osservate queste premesse il sistema potrà poi essere anche in grado di selezionare le copie più adatte ("appropriate") per un dato utente ("risoluzione selettiva").[10]

La soluzione illustrata da Caplan e Arms, in quanto basata su un sistema di identificatori univoci di opere,[11] dovrebbe comportare lo sviluppo di un meccanismo in grado di ottenere da ogni citazione un identificatore dell'opera citata, di norma con uno dei seguenti tre metodi:

1) estraendo l'identificatore incluso nella citazione (e ad esempio il DOI[12] spesso lo è);

2) calcolandolo in base ai dati bibliografici della citazione (alcuni identificatori sono predicibili e possono essere calcolati, ad es. il SICI);[13]

3) utilizzando i dati bibliografici della citazione per interrogare un apposito database esterno che fornisca il relativo identificatore.

A completare il sistema e garantire l'accesso alla copia più adatta per l'utente (quella per cui la sua istituzione o la sua biblioteca hanno un abbonamento attivo, quella collocata sul server più vicino o meglio collegato ecc.), lo studio

[6] Possiamo senz'altro affermare che a distanza di pochi anni questo "blocco culturale" è caduto, e anzi oggi i fornitori fanno vanto dell'approfondita e standardizzata raggiungibilità dei propri contenuti.

[7] Tra febbraio e giugno di quell'anno si sono svolti due workshop sponsorizzati dalla National Information Standards Organization (NISO), dalla Digital Library Federation (DLF), dalla National Federation of Abstracting and Indexing Services (NFAIS), dalla Society for Scholarly Publishing (SSP), e i lavori di un apposito gruppo; esito di queste attività sono stati due report e l'articolo a firma di Priscilla Caplan e William Y. Arms sopra citato.

[8] Ad esempio il PMID (PubMed Unique Identifier), identificatore numerico assegnato a tutti gli articoli censiti dal database PubMed. È dunque un identificatore "privato", ma la cui unicità è garantita all'interno del "sistema" PubMed.

[9] "Given the information in a standard citation, how does one get to the thing to which the citation refers?" (Priscilla Caplan – William Y. Arms, *Reference linking for journal articles*, cit.).

[10] La capacità di selezionare e fornire ad ogni utente la copia appropriata costituisce la soluzione del cosiddetto "Problema di Harvard", così descritto dal primo report del NISO Workshop: "In many cases there will be multiple copies of the same article available. For example, an Elsevier journal may be available in Science Direct, in Michigan's PEAK database, through OhioLink, etc. Many legitimate reasons for multiple copies exist, including performance (caching), different service models, archival needs, and competition. The system must be able to find the right copy for the user, which, in the end, may have more to do with who the user is than what the journal is".

[11] Per la verità Caplan e Arms affermano che, nonostante gli innegabili vantaggi dell'uso degli identificatori, è bene comunque trattarli come parte di un modello più generale, riconoscendo che vi possano essere casi in cui non si rendono necessari.

[12] Il DOI (Digital Object Identifier, <http://doi.org>) è un identificatore univoco di un qualsiasi oggetto di contenuto intellettuale, a qualsiasi livello di granularità (libro, capitolo, articolo, periodico, illustrazione, tabella…), in ambiente digitale, rispetto al quale è possibile identificare una proprietà intellettuale e associare dei metadati (essenzialmente informazioni bibliografiche, commerciali, di localizzazione). La risoluzione di un DOI, che si realizza tramite il protocollo Handle (di cui DOI è una più specifica applicazione), è teoricamente in grado di ridirigere l'utente dall'identificatore alle possibili localizzazioni (cioè URL) di copie dell'oggetto, ma poiché le attuali implementazioni (cfr. CrossRef, <http://www.crossref.org>) di norma registrano esclusivamente la localizzazione gestita dal fornitore del contenuto (tipicamente, quindi, la localizzazione sul sito dell'editore), non si possono considerare di per sé sistemi compiuti di "risoluzione selettiva". Il DOI è standard ANSI/NISO (Z39.84).

[13] Il SICI (Serial Item and Contribution Identifier) è un identificatore univoco di un volume, fascicolo, articolo, o altra parte identificabile di periodico. È un'estensione dell'ISSN (International Standard Serial Number) ed è standard ANSI/NISO (Z39.56).

propone l'integrazione del modello con un meccanismo di "risoluzione selettiva" dell'identificatore in una o più localizzazioni, cioè un meccanismo in grado di portare il lettore dall'identificatore alla copia o alle copie appropriate dell'opera, mediante l'interrogazione di un database di localizzazioni (o di due database a cascata, locale e globale), nonché di un database (il medesimo o altro) che contenga le informazioni necessarie per indirizzare l'utente alle copie cui ha diritto di accedere.
La possibilità di implementare, in alternativa all'estrazione di identificatori univoci e alla loro risoluzione selettiva, un sistema dinamico di costruzione di link computati "al volo" – solo nel momento in cui l'utente li percorra – sulla base dei metadati disponibili, viene citata come dotata di qualche vantaggio, ma considerata "probabilistica", dunque incapace di garantire il raggiungimento della copia desiderata, anche se esistente.
Il modello proposto, dunque, rimane nell'ambito delle soluzioni di linking statico, per quanto funzionalmente potenziato dalla sostituzione del link diretto alla copia con il sistema di risoluzione degli indicatori. A conclusione dello studio, tuttavia, gli autori ammettono che per gestire il linking citazionale entro universi documentari più ampi e complessi del solo *corpus* degli articoli scientifici potrebbe rendersi necessaria l'introduzione di tecniche di linking dinamico.
E difatti, all'incirca nello stesso periodo, Herbert Van de Sompel e Patrick Hochstenbach, che hanno avviato – dapprima presso l'Università di Gand, poi anche in collaborazione con il National Laboratory di Los Alamos – una sperimentazione congiunta basata su un approccio al problema più ampio e "visionario", ne espongono i primi risultati, cominciando a delineare la cornice concettuale di un sistema di linking aperto e dinamico.[14]
Van de Sompel assume che il problema del *reference linking* non sia solo un problema di fornitura della copia appropriata della risorsa citata, né tanto meno limitato alla specifica tipologia degli articoli di pubblicazioni periodiche, ma sia piuttosto una sottospecie di una questione più generale e inclusiva, e cioè: "Dati dei metadati bibliografici, come si possono presentare servizi estesi per essi?".[15]
Di conseguenza, delinea un'estensione delle potenziali funzioni del sistema, finalizzato a fornire non tanto o non solo una copia della risorsa citata, quanto piuttosto una serie di possibili servizi costruiti a partire dai dati bibliografici contenuti nella sua citazione. Per servizi estesi rispetto a una certa entità bibliografica si intende, ad esempio, il testo pieno dell'articolo citato in una bibliografia o in un database, un record di un altro database che ne contiene un abstract, le informazioni citazionali sull'articolo, sul suo autore, sulla rivista, le biblioteche che ne possiedono una copia cartacea, la recensione del libro citato, le librerie online che ne hanno una copia in catalogo ecc.
Rispetto ai tipici sistemi implementati dai fornitori commerciali (statici, chiusi e interni), Van de Sompel propone soluzioni architetturali diametralmente opposte, e cioè:
– un modello dinamico, in cui i link non vengono precalcolati e scritti "nella citazione", ma computati "al volo" solo al momento in cui servono per portare alla risorsa di destinazione;
– una struttura di linking aperta, in quanto consente che i link non vengano imposti dal servizio che fornisce la citazione, ma vengano scelti e forniti da una "terza parte", tipicamente l'istituzione dell'utente;
– un modello che integri tutte le collezioni, i database, i cataloghi, le risorse e i servizi d'interesse della biblioteca o del sistema bibliotecario, e non limitato ad una collezione i cui confini sono definiti da fornitori o aggregatori.
Rispetto a quello prefigurato dal rapporto di Caplan e Arms, Van de Sompel disegna dunque un modello più flessibile ed euristico, non necessariamente legato allo sviluppo di un sistema completo di identificatori (e quindi all'allestimento dei relativi database), in quanto prevede tecniche di risoluzione selettiva basate anche solo sull'uso dei metadati bibliografici standard (autori, titoli, date, pagine ecc.); e lo realizza con l'implementazione della cosiddetta "soluzione OpenURL" (o "schema OpenURL"), un'architettura che prevede una separazione fra la fornitura dei metadati (che deve avvenire da parte della risorsa di partenza, quella che contiene la citazione) e la fornitura dei servizi di linking, che avviene invece da parte di un software "interposto" (il "componente di servizio", il *link resolver*, cioè sostanzialmente un "risolutore di OpenURL"), tipicamente gestito dalla biblioteca o dalla sua istituzione, che intercetta i metadati, ed è configurato in modo da interpretarli e fornire in base ad essi uno o più link coerenti per l'intera collezione e sensibili al contesto dell'utente.[16]
Analizzando un po' più in dettaglio la dinamica realizzativa di questa architettura[17] possiamo così descriverne i passi principali:

[14] Il primo contributo di Herbert Van de Sompel – Patrick Hochstenbach, *Reference linking in a hybrid library environment, part 1: frameworks for linking*, che sviluppa il modello teorico del linking dinamico e aperto, viene pubblicato su "D-Lib Magazine" nell'aprile del 1999. Testo originale su <http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html>, trad. it. a cura di Cinzia Bucchioni su <http://www.aib.it/aib/commiss/cnur/trsomp1.htm3>.
[15] L'enunciazione è contenuta in Herbert Van de Sompel – Patrick Hochstenbach, *Reference linking in a hybrid library environment, part 3: generalizing the SFX solution in the "SFX@Ghent & SFX@LANL" experiment*, "D-Lib Magazine", 5 (1999), 10. Testo originale su <http://www.dlib.org/dlib/october99/van_de_sompel/10van_de_sompel.html>, trad. it. a cura di Cinzia Bucchioni su <http://www.aib.it/aib/commiss/cnur/trsomp3.htm3>.
[16] Una parziale convergenza delle due ipotesi (Caplan-Arms e Van de Sompel-Hochstenbach) si realizzerà mediante l'implementazione della conformità allo standard OpenURL da parte del database CrossRef, cioè della capacità di reindirizzare un identificatore DOI a un *resolver* istituzionale invece che alla copia del testo sul sito dell'editore.
[17] Ma naturalmente la nostra descrizione è comunque molto sintetica: per una più dettagliata esposizione, comprensiva dell'analisi degli aspetti più problematici dell'implementazione, si vedano le tre parti di *Reference linking in a hybrid library environment* sopra citate.

– la risorsa di partenza formatta metadati e identificatori relativi ad ogni citazione bibliografica secondo una sintassi standard, cioè la sintassi OpenURL, che tipicamente impacchetta i dati in una concatenazione di coppie nome/valore;[18]
– la risorsa di partenza conosce o è in grado di recepire "al volo" l'indirizzo del servizio di risoluzione di OpenURL dell'utente (cioè la URL del *link resolver*); questa "conoscenza" può essere realizzata in vari modi, uno dei quali, forse il più semplice e anch'esso ideato dal gruppo di Van de Sompel, è il cosiddetto "cookiepusher";[19]
– la risorsa di partenza presenta dunque, accanto ad ogni citazione, anche un "oggetto" che appare all'utente come una URL cliccabile (un pulsante o un link testuale), costruita in modo da inviare – al clic dell'utente – il pacchetto di metadati formattati al suo *link resolver* (o semplicemente *resolver*);
– il *resolver* riceve in input la OpenURL e in tal modo raccoglie metadati e identificatori (sia direttamente sia "andandoseli a prendere" altrove, se quelli ricevuti non sono sufficienti), ne fa una "valutazione contestuale" (cioè in base alla tipologia dei dati ricevuti e al contesto della collezione di cui l'istituzione dispone), dopo di che presenta all'utente un menu di link ai servizi pertinenti e sensibili al contesto;
– i link del menu di servizi vengono costruiti dal *resolver* utilizzando uno o più metadati e identificatori ricevuti dalla OpenURL, mediante una sintassi di puntamento specifica per ogni risorsa di arrivo.[20] Nelle scelte di costruzione del menu, nella selezione dei servizi da presentare, dei metadati da utilizzare per proporre percorsi di ricerca o di approfondimento, nell'organizzazione logica e funzionale delle opzioni, nell'integrazione con gli altri servizi della biblioteca (dalla ricerca a catalogo allo scarico della citazione, dal document delivery al reference digitale…) entrerà in gioco la capacità dell'istituzione che gestisce il *resolver* di innestare la conoscenza, l'esperienza, i percorsi e le procedure di un "bibliotecario di reference virtuale" al servizio dell'utente.

Questo, sia pure in sintesi, lo schema di funzionamento del modello proposto da Van de Sompel e il suo gruppo di collaboratori.[21] Ma in fondo, al di là delle questioni architetturali, che possono sembrare complicate ma sono di natura implementativa, ciò che forse stupisce di più di questo modello è la sua "naturalezza", proprio nel senso di capacità di ottenere nell'ambiente digitale e per via computazionale risultati analoghi a quanto otterrebbe un ricercatore umano

[18] Per un esempio di OpenURL conforme alla prima specifica, poi identificata come versione 0.1, vedi l'appendice 3. Si noti che col nome OpenURL, oltre allo schema organizzativo più generale, Van de Sompel battezza anche il "gancio", cioè il pacchetto di dati e identificatori codificati in base ad una specifica sintassi per il trasporto sulla rete e la trasformazione in link. Il successivo standard ANSI/NISO Z39.88 (*The OpenURL framework for context-sensitive services*, <http://www.niso.org/standards/standard_detail.cfm? std_id=783>), che nel 2004 codifica l'OpenURL 1.0, definisce e standardizza le componenti del sistema, cioè fondamentalmente la costruzione dei pacchetti di informazione e i metodi per trasportarli sulla rete. Il "pacchetto" di metadati in senso stretto viene denominato ContextObject (come già nell'implementazione del gruppo di Van de Sompel), e la sua sintassi diviene molto più complessa, comprensiva e generalizzata rispetto a quella dell'OpenURL 0.1. Per un primo approccio allo standard e una riflessione sulle principali differenze rispetto alla prima specifica si legga Cinzia Bucchioni, *L'evoluzione dello standard OpenURL*, "Bibliotime", 7 (2004), 3, <http://didattica.spbo.unibo.it/bibliotime/num-vii-3/bucchion.htm>.

[19] Per un'illustrazione del meccanismo *cookiepusher* si veda Herbert Van de Sompel – Patrick Hochstenbach, *Reference linking in a hybrid library environment, part 3*, cit. Sono attualmente ben consolidate, soprattutto nelle risorse dei fornitori commerciali, soluzioni "lato server" in cui la presentazione del link e l'indirizzamento al *resolver* appropriato sono completamente amministrati dal gestore della risorsa di partenza, cui l'istituzione che gestisce il *resolver* deve dunque sempre rivolgersi per l'attivazione. Si stanno però diffondendo e affermando anche soluzioni di indirizzamento al *resolver* "lato client", basate su COinS (ContextObjects in Spans), una specifica per l'incapsulamento dei metadati relativi all'oggetto citato nella pagina web, anche statica, che contiene la citazione. L'uso di COinS consente quindi a browser dotati di apposite estensioni o *bookmarklet* e adeguatamente configurati, di riconoscere la presenza di OpenURL latenti nelle pagine web e presentare il link indirizzato al *resolver* appropriato per l'utente. Interessanti anche le soluzioni basate sull'interfacciamento con *router* di indirizzamento nazionali (si veda l'Information Environment Service Registry, IESR, di JISC, all'indirizzo: <http://iesr.ac.uk/>) o internazionali (si veda l'OpenURL Resolver Registry di NISO, attualmente implementato e mantenuto sperimentalmente, all'indirizzo: <http://www.oclc.org/productworks/urlresolver.htm>). Mentre nel caso del gateway IESR di JISC si tratta di un servizio integrato nella struttura della rete nazionale e di uso esteso, nel secondo caso si ha un servizio più che altro dimostrativo, che consta di un repertorio di *resolver* ad adesione volontaria, in cui è possibile descrivere le conformità tecniche del proprio sistema e associarvi un range di IP (tipicamente quello dell'istituzione), a supporto del gateway di reindirizzamento: pertanto, inviando un'OpenURL al gateway da un IP associato a un *resolver* registrato, avviene automaticamente il reindirizzamento dell'OpenURL al *resolver* associato.

[20] Per la precisione, nell'implementazione di Van de Sompel i link non vengono risolti in URL al momento della presentazione del menu di servizi, ma solo nel momento in cui l'utente li percorre effettivamente. È solo in quel momento, infatti, che un apposito modulo software costruisce "al volo" – utilizzando metadati e identificatori disponibili – l'indirizzo cui puntare, implementando la sintassi di linking specifica della risorsa di arrivo. Così come la sintassi, anche la precisione e la profondità del linking possono variare risorsa per risorsa: ciò vuol dire che, a seconda dei metadati disponibili e della capacità della risorsa di arrivo di essere linkata in profondità, si può andare dal puntamento diretto al full text di un articolo mediante una URL stabile basata su un identificatore, al puntamento a pagine di TOC o di testate, fino a soluzioni che lanciano query basate su metadati bibliografici che possono dare risultati di diversa accuratezza e granularità.

[21] Modello che naturalmente ha negli anni successivi conosciuto evoluzioni, standardizzazioni, integrazioni, varianti architetturali e implementative, estensioni, revisioni critiche, e ancora nuove sperimentazioni, applicazioni, e così via. Rendere conto di tutto ciò richiederebbe uno spazio che va molto al di là degli obiettivi di questo articolo, e potrà magari essere oggetto di un contributo successivo.

in un ambiente documentario tradizionale, simulando molto da vicino i suoi comportamenti, le sue euristiche, e rinunciando volutamente ad alcune delle formalizzazioni tipiche dell'ambiente elettronico, come appunto il vincolo di identificazione univoca degli oggetti, a favore di una maggiore flessibilità.
Eppure basta pensare a come abbiamo sempre maneggiato nella nostra vita di lettori o di bibliotecari le citazioni bibliografiche, non sempre e solo per recuperarne l'oggetto, ma anche per partire per altri percorsi, altri testi dello stesso autore o informazioni su di lui, altri saggi in quel fascicolo monografico, altri eventi di quel convegno, ciò che ha scritto dopo quel coautore, l'impact factor di quella rivista, i libri e gli articoli che citano quell'articolo o quel libro, il suo abstract o le sue recensioni, e così via, per capire che forse è attraverso questa naturalezza che si supporta più efficacemente nell'ambiente digitale il comportamento associativo della mente umana.

1.3. Per concludere

Per concludere questa riflessione introduttiva al contesto, agli approcci e alle motivazioni del linking citazionale, torniamo ad assumere il "punto di vista" delle biblioteche e dei bibliotecari, e a focalizzare alcuni vantaggi dell'implementazione di questo modello di *reference linking*:
– fornisce agli utenti un sistema di navigazione coerente, teso alla minimizzazione dei vicoli ciechi, fra le risorse distribuite che costituiscono la collezione digitale dell'istituzione bibliotecaria;
– consente di integrare e valorizzare nella collezione digitale anche tipologie di risorse che spesso ne restano ai margini, sono poco conosciute dagli utenti, come ad esempio le iniziative di editoria *free-online*, gli archivi open access, e in generale tutte le risorse che non sono collocate su piattaforme commerciali "forti", e che per questo rischiano di sparire dall'orizzonte conoscitivo dell'utente indipendentemente dalla loro rilevanza scientifica;
– valorizza gli strumenti e i servizi "core" della biblioteca: i cataloghi (sia nel ruolo tipico di localizzatori delle risorse sia in quello innovativo di base di conoscenza del posseduto elettronico), e quindi i servizi di prestito e di consultazione; i servizi di prestito interbibliotecario e di fornitura di articoli (che fungono da "integratori" fra collezione digitale e cartacea);
– consente di fare *branding* e fidelizzare gli utenti (anche quelli remoti) alla biblioteca, al sistema bibliotecario e all'istituzione, piuttosto che alle piattaforme commerciali: la presenza di un'icona che richiama il logo della biblioteca o del sistema anche nelle pagine delle risorse commerciali, la presentazione di un menu di servizi (che può essere anch'esso evocativo dell'istituzione nei colori, nel titolo, nel logo ecc.) che funge da "connettivo" fra le risorse e da veicolo verso ulteriori servizi, non possono che consolidare l'identità, la riconoscibilità, il valore simbolico dell'istituzione bibliotecaria;
– restituisce alle biblioteche il loro naturale ruolo di organizzazione, gestione e integrazione delle informazioni e dei servizi, e di mediazione per l'accesso alla conoscenza, tanto più importante in un contesto in cui la fruizione delle risorse delle biblioteche si sposta sempre di più al di fuori delle loro mura.
Dunque, riprendendo la metafora conversazionale del *Manifesto per le biblioteche digitali* citata in apertura, fra i tanti strumenti per l'interoperabilità che potranno supportare le "conversazioni" tra biblioteche, e tra biblioteche, bibliotecari e utenti, l'OpenURL si candida a diventare uno dei linguaggi principali, anzi, come ha già detto qualcuno, "l'esperanto del linking".[22]

2. L'OpenURL *resolver* dell'Università di Pisa

In questa seconda parte del contributo intendiamo presentare le fasi di sviluppo e le specifiche funzionali dell'OpenURL *resolver* da noi sviluppato e attualmente in uso all'Università di Pisa.
Il progetto ha preso avvio da percorsi di studio individuali sull'ideazione e lo sviluppo della soluzione di *reference linking* contestuale "standard OpenURL + OpenURL *resolver*", documentati nella già citata serie di articoli su "D-Lib Magazin" tra il 1999 e il 2001:[23] tale soluzione ci aveva incuriosito, oltre che per le funzionalità promesse[24] (e promosse in nuovi prodotti commerciali),[25] per la sua genesi

[22] La frase "OpenURL is the esperanto of linking" è di Chuck Koscher, technology director di CrossRef.
[23] Si vedano i già citati articoli di Herbert Van De Sompel et al., pubblicati su "D-Lib Magazine" fra il 1999 e il 2001, e le traduzioni di essi disponibili sul sito della Commissione nazionale università e ricerca dell'Associazione italiana biblioteche a partire dallo *Scaffale CNUR*, <http://www.aib.it/aib/commiss/cnur/scaffale.htm3>.
[24] Per approfondire la tecnologia in questione la documentazione in rete è ormai vasta: si vedano almeno in italiano i due articoli di Cinzia Bucchioni, *SFX e OpenURL: gli esperimenti del team di Van de Sompel*, "Bibliotime", 5 (2002), 2, <http://www.spbo.unibo.it/bibliotime/num-v-2/bucchion.htm> e *L'evoluzione dello standard OpenURL*, cit.; inoltre, almeno: Ann Apps, *OpenURL and OpenURL framework: demystifying link resolution*, "Ariadne", 38 (gennaio 2004), (*Report on conference about OpenURL held by NISO in Washington*), <http://www.ariadne.ac.uk/issue38/apps-rpt/intro.html>. Per il testo dello standard si veda il sito: <http://www.niso.org/standards/standard_detail.cfm?std_id=783>.
[25] Sulle sperimentazioni del team di Van de Sompel aveva richiamato la nostra attenzione Antonella De Robbio in una chiacchierata oltremodo stimolante, da cui tra l'altro era emersa la necessità di documentarsi direttamente sulle tecnologie alla base dei nuovi prodotti, evitando di affidarsi esclusivamente alle promozioni commerciali; necessità tanto ovvia quanto sempre più trascurata, anche solo per la difficoltà di tenere il passo con le accelerazioni del mercato; ma proprio l'*overflow* delle proposte commerciali, che non trova corrispondenza in un'adeguata disponibilità di spesa, impone scelte e dunque strategie di orientamento autonomo: in questo senso la confusione tra aggiornamento professionale e informazione commerciale, da tempo ben conosciuta in altri settori a forte innovazione tecnologica, appare un rischio anche per il mondo bibliotecario.

applicativa in risposta ad esigenze concrete e specifiche,[26] poi approdata alla proposta di uno standard di settore.
Proprio il carattere eminentemente pratico della soluzione ha reso ben presto evidente l'impossibilità di valutarne a pieno funzionamento, complessità, criticità senza "provarla": abbiamo così installato su un server web, e messo a punto, il software dimostrativo di Andy Powell,[27] uno *script* Perl molto sintetico e semplice, che sostanzialmente imposta la ricezione e l'analisi dei parametri di una OpenURL prima generazione, e l'apertura di una schermata di servizi generici.
Subito dopo, trattandosi di un dispositivo di interoperabilità per la produzione di servizi di rete, è sorta l'esigenza di valutare sia la tenuta dello standard nel colloquio con sistemi "altri", sia l'impatto del servizio sui vari servizi della biblioteca e sugli utenti "reali". Era dunque necessario inserire il *resolver* nel contesto di una vera "biblioteca ibrida": abbiamo perciò proposto la sperimentazione al Servizio per il Sistema bibliotecario dell'Università di Pisa, presso il quale abbiamo trovato accoglienza entusiasta e generosa e ampia collaborazione.

2.1. La filosofia di sviluppo

La nostra filosofia di sviluppo si è mossa su due binari: uno fortemente pragmatico, orientato allo sviluppo di servizi immediatamente utili nel contesto; l'altro indirizzato a tenere comunque d'occhio le funzionalità basilari e irrinunciabili di un OpenURL *resolver*.
Il primo binario muove dalla nostra interpretazione di questa tecnologia come strumento eminentemente di intermediazione bibliotecaria, cioè mezzo per offrire al bibliotecario la possibilità di intervenire nei percorsi di ricerca dell'utente, approntare scorciatoie, introdursi (introdurre la propria consulenza professionale) nelle interfacce eterogenee dei vari fornitori e dei vari ambienti della biblioteca digitale, riconducendole all'unità delle finalità bibliotecarie e ove necessario al contesto della biblioteca ibrida. In tale ottica, questa tecnologia ha senso solo se assolutamente "personalizzata" rispetto alla biblioteca specifica (contestualizzata): la configurazione del sistema, i servizi attivati e il loro ordine di presentazione, le modalità e il livello comunicativo devono riflettere la situazione locale; necessitano dunque di una decisa selezione dei percorsi da proporre e di una coraggiosa scelta e gerarchizzazione delle risorse su cui si ritiene opportuno indirizzare gli utenti per l'uso ottimale della biblioteca.
Il secondo binario, mentre tiene fede agli iniziali interessi teorici, risponde alla nostra convinzione che una seria fondazione standard sia comunque imprescindibile, non solo per eventuali generalizzazioni/esportazioni del nostro *resolver*, ma anche per una sua speranza di vita nel contesto di partenza, il quale è ovviamente (come tutto) in rapida evoluzione.
Sulla base di questi presupposti, progressivamente e parallelamente abbiamo proceduto a implementare nello *script* la risoluzione della OpenURL 1.0, e ad espandere il *range* delle *source* e dei *target* del nostro *resolver*: per pura semplicità espositiva i vari passi vengono presentati in successione.

2.2. La nuova OpenURL

Porre il nostro *resolver* in linea con lo sviluppo dello standard[28] è stata la prima questione che ci si è posta, essendo lo *script* di partenza impostato per la sola OpenURL di prima generazione (versione 0.1). L'interpretazione delle OpenURL versione 1.0 (o standard NISO Z39.88)[29] è risultata abbastanza semplice per quanto riguarda la conformi-

[26] Altro aspetto problematico del rapporto con le tecnologie (in verità diversa declinazione del precedente) è il rischio di uno sviluppo autoreferenziale di queste ultime: se continuiamo a ritenere la tecnologia uno strumento per risolvere problemi e non un fine, allora bisogna dedurre che non sempre l'ultima e massima tecnologia possibile è la miglior soluzione in ogni situazione – e a chi gestisce servizi bibliotecari è necessaria l'analisi dei bisogni (compresi quelli inespressi o impliciti, anche indiretti) prima che delle tecnologie possibili.
[27] Una presentazione completa dello *script* nell'articolo di Andy Powell, *OpenResolver: a simple OpenURL resolver*, "Ariadne", 28 (giugno 2001), <http://www.ariadne.ac.uk/issue28/resolver/>.
[28] Per una piena comprensione delle problematiche dell'allineamento alla nuova versione è ovviamente necessario avere presenti i dettagli tecnici dello standard; si raccomanda pertanto la lettura del testo dello standard ANSI/NISO Z39.88-2004 (*The OpenURL framework for context-sensitive services*, cit.); un ausilio a tale lettura può essere il già citato articolo di Cinzia Bucchioni, *L'evoluzione dello standard OpenURL*, cit.
[29] La complessità del nuovo standard non può essere esaurita in una nota, proviamo comunque a ricordarne brevemente i punti qualificanti: esso ha carattere di *framework* aperto ed estensibile, basato su un "registro online" (<http://alcme.oclc.org/openurl/servlet/OAIHandler?verb=ListSets>) in cui sono definiti (ed è sempre possibile aggiungerne di nuovi) formati, identificatori, metadati ecc., da utilizzarsi poi nelle applicazioni; questo perché, rispetto all'impostazione iniziale, la nuova versione intende generalizzare per potersi applicare anche a contesti diversi dall'universo bibliografico; definisce dunque un ContextObject, cioè un costrutto informativo atto a viaggiare sulla rete, che contiene la descrizione di un oggetto citato sulla rete (il *referent*: la pubblicazione, ma eventualmente anche una persona, un'azienda, un'immagine) e il "contesto di rete" in cui tale oggetto viene citato (per esempio, servizio di provenienza dei dati, utente richiedente, servizio richiesto ecc.); le applicazioni basate sullo standard possono quindi potenzialmente svilupparsi in diversi domini (per esempio non solo bibliografico ma museale, cinematografico ecc.) e per servizi anche lontani dall'iniziale *reference linking*. Per sviluppare le applicazioni concrete è pertanto necessario individuare diversi "profili di comunità" o sottoinsiemi dello standard, ciascuno dei quali riferito a determinati oggetti e servizi e comprendente determinati formati, metadati, identificatori; il Profilo SAP (*San Antonio Profile*) è il sottoinsieme dello standard selezionato come rilevante per la gestione dei servizi bibliografico-accademici: pertanto nel nostro caso conformità con lo standard Z39.88 ha significato conformità con il Profilo SAP, nel cui ambito sono definiti due livelli, SAP1 e SAP2, i quali si distinguono per una diversa modalità di formalizzazione: SAP1 struttura i metadati in coppie nome/valore (proprio come la vecchia OpenURL), mentre SAP2 prevede l'uso di schemi xml. Il nostro sviluppo si è fatto carico della conformità a SAP1; per la conformità al profilo SAP2 è stata svolta l'analisi preliminare, e si ritiene che l'interpretazione di tali OpenURL

tà al *community profile* SAP1 (cioè il profilo specifico per la letteratura scientifico-accademica, che ripropone con la nuova sintassi i metadati e i concetti della OpenURL 0.1): a questo fine abbiamo aggiunto al *resolver* la "conoscenza" di metadati e identificatori[30] espressi nella sintassi 1.0, e (più complesso) delle modalità di espressione *by value* e *by reference* sia per i metadati che per il ContextObject. In questa fase ci siamo attenuti alle linee guida per l'upgrade e il controllo di conformità messe a disposizione sul sito dello standard e sui siti collegati.[31]
È evidente che un ben diverso grado di complessità si genererebbe nel programma quando si volessero prendere in conto altri profili, ma soprattutto quando ci si volesse far carico dello standard nella sua generalità: l'estensibilità del "registro online"[32] a nuovi formati, identificatori, metadati renderebbe desiderabile l'implementazione di un colloquio automatico registro/*resolver* che ponesse il *resolver* in grado di consultare direttamente il registro per trarne gli elementi nuovi: tale sviluppo, che peraltro va oltre i requisiti dello standard e non è neppure ben supportato dal registro nella sua forma attuale (non tutte le definizioni sono in formato leggibile dalla macchina), pare per ora tralasciato dalle implementazioni concrete.[33]

2.3. Le source

Source sono definite, in base alla terminologia OpenURL, le risorse di partenza, quelle cioè che aderiscono allo standard e dunque sono in grado di inviare una OpenURL a un *resolver*; in altre parole, le risorse che, se del caso, sono in grado di presentare in corrispondenza delle citazioni bibliografiche, un link al *resolver*, cioè un bottoncino che l'utente può "cliccare" facendo aprire la schermata del *resolver*.[34]
Attivare le *source* ha dunque significato verificare quali, tra le risorse digitali dell'ateneo, sono OpenURL *compliant* e comunicare loro l'esistenza di un *resolver* istituzionale. Semplice dunque: questo è il versante regolamentato dallo standard, e avendolo la risorsa *source* implementato, il colloquio è immediato.
Le modalità pratiche sono risultate comunque piuttosto varie:
– per i fornitori che gestiscono centralmente i profili utente è stata sufficiente una mail di richiesta con comunicazione dell'indirizzo di rete di *resolver* e icona;
– per i fornitori che lasciano al cliente possibilità di intervento sul profilo utente si è provveduto a configurare le impostazioni necessarie sul sito remoto, in alcuni casi direttamente costruendo la OpenURL a partire da selezioni di campi del database;
– per alcune risorse settoriali è stato infine necessario fornire previa dimostrazione del funzionamento del sistema.[35]
Attivato il link, dunque il colloquio risorsa/*resolver*, abbiamo constatato – come peraltro previsto – l'esistenza di varie idiosincrasie nell'implementazione dello standard. Per quasi ogni *source* è risultato necessario prevedere nel *resolver* trattamenti specifici per le rispettive OpenURL: o per l'utilizzo di dati privati (previsti se pur sconsigliati nello standard); o per metadati, o sintassi, devianti; o anche solo per migliorare le prestazioni (per esempio nei casi in cui il metadato "title" conteneva anche sottotitoli e altri elementi della prima area ISBD).

2.4. L'arricchimento dei metadati

Abbiamo poi riscontrato, da parte di alcuni servizi, l'invio di OpenURL con metadati molto poveri, oppure con soli identificatori. In entrambi i casi, abbiamo ritenuto opportuno implementare per quanto possibile il cosiddetto "arricchimento dei metadati", cioè fare in modo che il *resolver*, prima di presentare il menu all'utente, provi a procurarsi metadati descrittivi migliori.
Il secondo caso merita una riflessione: se, come già detto, proprio degli identificatori è identificare univocamente una risorsa, la loro efficacia in rete sta nel poter condurre direttamente alla risorsa stessa: dunque l'utilizzo più immediato

potrà avvenire senza problemi non appena si presentino applicazioni *source* che inviino metadati in tale modalità.
[30] Restano da implementare alcuni identificatori poco utilizzati nelle nostre collezioni, per cui non è stato ancora possibile operare test (per esempio OCLCNum). Inoltre ci siamo limitate a implementare la risoluzione dei metadati che descrivono il *referent*, e le entità contestuali strettamente necessarie per i servizi previsti ed effettivamente presenti nelle OpenURL ricevute dai servizi reali.
[31] *Description of San Antonio (level 1, level 2). Community profiles*, <http://alcme.oclc.org/openurl/docs/pdf/SanAntonioProfile.pdf>; *The Key/Encoded-Value (KEV) format-implementation guidelines*, <http://alcme.oclc.org/openurl/docs/implementation_guidelines/KEV_Guidelines-20040211.pdf>; *OpenURL 1.0 test suite*, <http://swets.1cate.com/t/test/niso-suite-10.html>.
[32] Il registro, che risulta elemento chiave dello standard, è affidato ad una *maintenance agency* che attualmente è OCLC.
[33] Se la fortuna dello standard, nella sua prima versione, è stata decretata proprio dall'immediata applicabilità e dall'estrema semplicità, aspetti che hanno incoraggiato l'ampia e pronta adesione da parte dei fornitori commerciali (l'investimento necessario per rendersi conforme essendo bassissimo), è fondato il timore che la formalizzazione, generalizzazione e conseguente complessità della nuova versione dello standard, invece che innescare l'auspicata estensione al di là del dominio e dei servizi di partenza, possano piuttosto diventare un ostacolo al suo sviluppo; tutti gli ausili messi a disposizione dagli sviluppatori per aiutare l'upgrade dei *resolver* ma anche delle *source*, manifestano consapevolezza del rischio. Peraltro attualmente le applicazioni al di fuori del *reference linking* bibliografico sono solo sperimentali e restano comunque legate al mondo della documentazione.
[34] Ripetiamo: il clic dell'utente chiama il *resolver* istituzionale, passandogli i dati bibliografici della citazione, opportunamente formattati secondo le specifiche dello standard OpenURL; il *resolver* è così in grado di costruire una schermata ad hoc, con un menu di servizi (*link*) selezionati ed impostati opportunamente di volta in volta.
[35] PubMed, per esempio, prima di procedere a inserire il link a un *resolver* istituzionale lo testa relativamente alla risoluzione delle proprie OpenURL, le quali contengono descrizioni con solo identificatore PMID, dunque "funzionano" con *resolver* che implementino l'opportuno meccanismo di interpretazione.

degli identificatori sta nel prenderli come base del servizio di link[36] e non necessita di metadati né di arricchimenti.
Ma questo – che risulta comunque riduttivo rispetto alla complessità degli oggetti bibliografici, come già evidenziato dalle problematiche della copia appropriata – vale in ambiente pienamente elettronico, mentre il nostro contesto di riferimento, in questo caso non solo istituzionale ma in generale italiano, ci pare ancora fortemente ibrido, il che ci ha consigliato di sfruttare del *resolver* le potenzialità di "collegamento" tra le collezioni digitali e tradizionali, di rendere manifesti i rapporti di complementarietà tra le due; ovviamente le collezioni cartacee sono integrabili solo in quanto rappresentate dagli OPAC, cioè da cataloghi generati da prassi catalografiche tradizionalmente poco sensibili agli identificatori (scarsa attenzione a riportare i diversi ISBN di un'edizione; poca solerzia nell'integrare le descrizioni con ISSN assegnati in momenti successivi ecc.).
Alla luce di queste considerazioni, abbiamo deciso di lavorare il più possibile anche con i metadati e non con i soli identificatori, utilizzando questi ultimi per il reperimento dei metadati oltre che per l'indirizzamento di rete.
Gli arricchimenti implementati nel nostro *resolver* sono i seguenti:
– dal DOI ai metadati;
– dai metadati, se sufficienti, al DOI degli articoli;
– dall'ISSN al titolo del periodico (quando possibile);
– dall'ISBN al titolo della monografia (quando possibile);
– dal PMID[37] ai metadati;
– dall'identificatore OAI ai metadati.
Questa scelta è stata condizionata dai servizi di rete attualmente disponibili, e le relative procedure sono specifiche all'attuale configurazione di tali servizi, dunque a stabilità relativa; abbiamo verificato direttamente la lamentata mancanza, su questo versante, di una standardizzazione e di una generalizzazione[38] che possano garantire una ragionevole automazione della risoluzione (a parte l'eccezione del DOI, peraltro a gestione "privata");[39] per esempio mancano servizi di rete ufficiali completi, accessibili, tramite cui sciogliere ISSN[40] e ISBN.[41]

2.5. I target: servizi e tecniche

Target sono definite, in base alla terminologia OpenURL, le risorse di arrivo, quelle cioè cui il *resolver* reindirizza, data una certa citazione di partenza; in altre parole, i link (le risorse "linkate") presentati di volta in volta nel menu del *resolver*.
Attivare i *target* ha dunque significato:
– selezionare i link (o servizi) da proporre per ogni categoria di citazione;
– verificare le condizioni tecniche di attivazione di tali link, dato che lo standard OpenURL non si fa carico di questo aspetto.
L'elemento della selezione è quello che più direttamente dà forma alla "politica di reference" che il bibliotecario intende "incorporare" nel *resolver*; è dunque a questo punto che si rende necessaria l'analisi del contesto più dettagliata e approfondita onde valorizzare le risorse e i servizi localmente disponibili.
Da parte nostra, questo ha significato prendere atto della situazione bibliotecaria in cui il nostro strumento veniva ad inserirsi, caratterizzato da una situazione bibliografica altamente ibrida, dove alcuni settori disciplinari si sono significativamente spostati verso il digitale mentre altri restano in larga prevalenza cartacei; dove inoltre alcune classi (anche generazionali) di utenti vanno educate all'uso esteso del digitale, onde evitare il sottoutilizzo dei cospicui investimenti finanziari relativi, mentre altre vanno ricondotte all'uso del materiale cartaceo quando necessario, onde evitare il sottoutilizzo del patrimonio bibliografico sedimenta-

[36] Per esempio usarli come indirizzo (o meglio: per costruire l'indirizzo) di rete del "testo pieno": il DOI può condurre al testo dell'articolo tramite il DOI *resolver*, l'ISSN, con l'ausilio di opportune liste, può essere usato per l'accesso al sito del periodico elettronico.
[37] Cfr. nota n. 8.
[38] Si ricordi sull'argomento almeno l'illuminante relazione di Juha Hakala, *Principi di identificazione: prospettive europee*, presentata alla "International Conference Electronic Resources Definition, Selection and Cataloguing", tenutasi a Roma nel novembre 2001; <http://w3.uniroma1.it/ssab/er/relazioni/hakala_ita.pdf>.
[39] Il sistema DOI è comunque molto ben integrato nella struttura del web: si basa su un sistema di server che (limitatamente agli editori che aderiscono a titolo volontario) conduce direttamente dal codice ai metadati, o dal codice alla localizzazione del testo pieno. Attraverso l'agenzia CrossRef (http://www.crossref.org/) è possibile per le biblioteche accedere a tali informazioni previa iscrizione gratuita: ci siamo avvalse di questa opportunità.
[40] Per ISSN esiste invero l'agenzia ufficiale (per il solo reperimento dei metadati corrispondenti: <http://www.issn.org/>), in questo caso il problema è di accessibilità, in quanto l'accesso richiede una sottoscrizione onerosa; per questo, anche se abbiamo implementato l'accesso tramite lo *ISSN register* (utilizzando l'accesso di prova), il funzionamento a regime dell'istallazione pisana si limita all'arricchimento sul catalogo nazionale ACNP, che per quanto ovviamente meno completo consente libero accesso e semplicità di prelevamento dei metadati, in formato XML.
[41] Per gli ISBN lo scioglimento (intendendo anche qui il reperimento dei metadati corrispondenti) è in linea di principio garantito dalle bibliografie nazionali: è evidente quanto poco questo sia utilizzabile dal punto di vista implementativo. Giusto per fare il caso italiano, considerando l'OPAC SBN il miglior candidato per la risoluzione degli ISBN italiani (non essendo BNI un servizio web), piuttosto grave risulta l'impossibilità di trovarvi un ISBN se nella stringa di ricerca mancano i trattini: molti servizi infatti passano ISBN senza trattini, e mentre è sempre possibile toglierli, non altrettanto è facile introdurli. Di fatto la nostra scelta, molto pragmatica, è stata di reperire i metadati mediante accesso al catalogo collettivo britannico COPAC, il quale combina facilità e velocità di accesso e ricchezza per l'ambito accademico. Anche per lo scioglimento degli identificatori OAI (Open Archive Initiative) dei documenti depositati negli archivi aperti pare mancare per ora un servizio ufficiale, e soprattutto stabilmente accessibile, e sono state perciò impostate risoluzioni su siti disponibili ma poco stabili.

tosi nel tempo, per il quale non si può contare su una massiccia riconversione al digitale, o per mancanza di investimenti editoriali nella riconversione[42] o per il mancato acquisto del nuovo supporto da parte della biblioteca. La selezione e la presentazione dei *target* è dunque stato per noi il modo per tradurre in menu di link quelli che sarebbero i suggerimenti del bibliotecario di reference di fronte a una citazione data.

Un altro aspetto del contesto locale che abbiamo individuato come problematico e passibile di impatto positivo, se pur non di soluzione, da parte del *resolver*, è la molteplicità di strumenti catalografici e repertoriali, generatasi per sedimentazioni successive a causa della varietà dei supporti documentali, delle eterogenee modalità di acquisizione, gestione, accesso ad essi, della instabilità di alcune collezioni (per esempio quelle editoriali di e-journal).

In una situazione in cui coesistono catalogo dei libri, catalogo (collettivo) dei periodici cartacei, lista dei periodici elettronici in abbonamento (magari una lista di ateneo affiancata da liste di struttura), lista dei database acquisiti, pagine web disciplinari con elenchi di siti di interesse (più o meno aggiornate) ecc., il *resolver*, presentando per ogni documento in un'unica schermata i link alle risorse (anche catalografiche) pertinenti, è un potente strumento di "rimessa in ordine" degli accessi alla ricerca.[43]

In considerazione della situazione, anche catalografica, descritta, è risultato opportuno funzionalizzare il menu presentato soprattutto al genere del documento: attualmente sono ben definiti e distinti i percorsi per monografie, articoli, periodici (sottoinsieme del precedente), tesi; per ora nella digital library di riferimento non abbiamo incontrato esempi concreti degli altri documenti previsti dallo standard. Per articoli e periodici il *resolver* lavora quando possibile con gli identificatori ISSN e DOI, sia per reperire i metadati sia per accedere ai testi. In assenza di ISSN, vengono proposte comunque le medesime ricerche tramite le parole del titolo, simulando – crediamo – il comportamento di un ricercatore umano.

Per quanto riguarda le monografie, non abbiamo ritenuto opportuno privilegiare la ricerca per ISBN pur in presenza di esso, in considerazione della possibile molteplicità di ISBN per edizione,[44] delle frequenti imprecisioni dei cataloghi, del darsi di casi in cui risulta utile il reperimento di altra edizione dell'opera citata. I servizi del *resolver*, basati essenzialmente su ricerche per titolo,[45] non risultano dunque particolarmente ridotti in assenza di ISBN nella OpenURL.[46]

Dal punto di vista implementativo, come accennato, l'accesso ai *target* non rientra nell'ambito dello standard OpenURL, né di altro standard specifico; questo significa che per ogni risorsa *target* da "linkare" bisogna riconoscere la specifica sintassi di link e implementarla,[47] quando possibile e al livello consentito (è possibile "linkare" la home del periodico, o anche ogni singolo fascicolo, o magari ogni singolo articolo?). Ciò nonostante, le caratteristiche del protocollo HTTP fanno sì che l'accesso ai vari *target*, seppur differente per ciascuno, risulti di solito di facile implementazione, almeno quando non siano posti in atto appositi meccanismi di autoprotezione da parte del sito cui si accede.

[42] È ben noto (ma non abbastanza ricordato) che per le lingue europee la conversione al digitale dei contenuti del patrimonio culturale procede con grande lentezza; ma anche per gli studi critici (le cosiddette fonti secondarie) continua ad esistere una produzione scientifico-accademica solo cartacea. Per inciso: i bibliotecari come mediatori documentali non possono non avere coscienza di quanta parte della produzione scientifico-editoriale italiana, ma anche europea, rimanga ai margini dei grandi circuiti produttivi e distributivi delle multinazionali editoriali, di conseguenza tagliata fuori non solo dal digitale ma anche dai principali database bibliografici.

[43] Continuiamo a ritenere desiderabile uno strumento di accesso unificato (il catalogo omnicomprensivo) alle risorse documentali della biblioteca (anzi il concetto di *one-stop-shop* perseguito da molti servizi web ci pare una sorta di "estensione" di tale classica aspirazione bibliotecaria); e sebbene forse non più praticabile come unitario sistema di controllo catalografico, vale certo la pena di provare a recuperare una "ricerca unificata" con nuovi strumenti. I già citati sistemi di *metasearch* rispondono esattamente a questa esigenza. Il *resolver*, pur non essendo questo il suo "core service", può dare una mano in tale direzione, per esempio se corredato di un'interfaccia di ricerca che sia *OpenURL generator*: l'utente potrà così immettere direttamente la sua stringa di ricerca in una normale *form*, che però invece di lanciare la ricerca su un catalogo genererà una OpenURL e una chiamata al *resolver*, il quale dunque reindirizzerà l'utente alle diverse risorse pertinenti, seppur non in un unico passo; nell'istallazione pisana è stato molto presto aggiunta questa funzionalità, su iniziativa dello SBA.

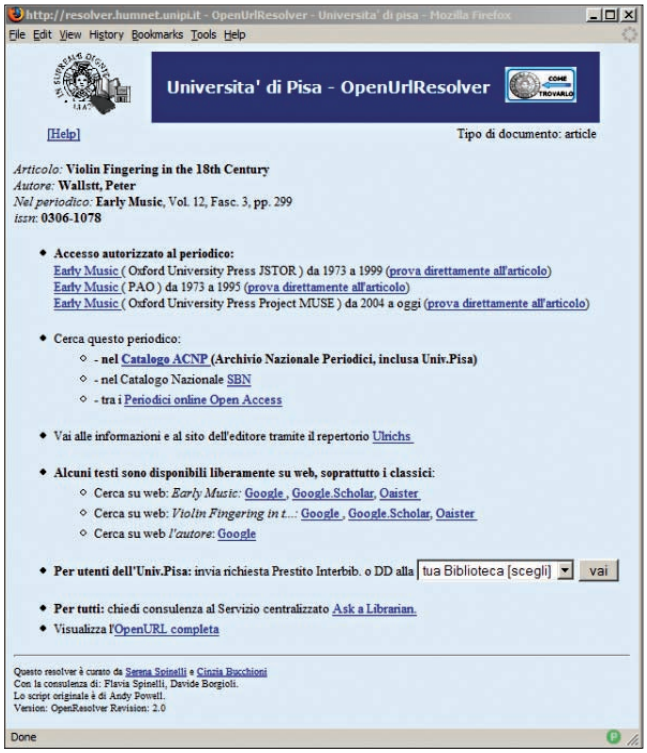
[44] Abbiamo indagato la possibilità di utilizzare il servizio xISBN di OCLC interrogandone il relativo database, che – dato un ISBN – dà in risposta l'elenco degli ISBN corrispondenti alla medesima edizione (<http://www.oclc.org/research/projects/xisbn/>), e abbiamo verificato che è facilmente integrabile nel *resolver*; ma per ora non abbiamo ritenuto opportuno inserirlo, trattandosi di servizio provvisorio su banca dati dimostrativa (e ovviamente non inclusiva della maggior parte delle pubblicazioni italiane).

[45] L'unico servizio che non viene proposto in assenza di ISBN è l'accesso ai siti del commercio elettronico, considerato nel contesto servizio secondario. Sui cataloghi è sempre proposta la ricerca per titolo, e in presenza di ISBN viene aggiunta anche la proposta di ricerca per ISBN.

[46] Assenza che può essere determinata tanto da incompletezza della OpenURL quanto da inesistenza dell'identificatore per un dato documento: come già accennato (cfr. nota 42), la grande massa di materiale bibliografico rappresentato negli OPAC precede l'era digitale e spesso anche l'era degli identificatori, né può essere archiviata *tout court* come irrilevante per la ricerca corrente, almeno non nell'ambito umanistico, in cui gli studi possono mantenere longeva validità e non per questo essere riconvertiti in digitale. Inoltre continua ad esistere in Italia una produzione accademico-editoriale addirittura priva di identificatori, una produzione tanto marginale rispetto ai circuiti produttivi e distributivi da correre il rischio di scivolare verso l'ambito della letteratura grigia (senza che peraltro venga approntato alcuno strumento di controllo ad hoc: non sempre un'università ha modo di render conto sistematicamente della propria produzione).

[47] Si tratta di brani di codice specifici per *target*, più o meno complessi, spesso di una sola riga, che costituiscono i cosiddetti *target-parser*.

Fig.1 - Esempio di menu di servizi per un articolo di periodico



L'accesso agli OPAC (semplice in quanto appunto risorse a libero accesso) è stato lasciato a livello di ricerca lanciata, il che significa che l'utente riceve una risposta, negativa o positiva, da vagliare.

L'accesso ai testi pieni degli articoli di periodico viene mediato dalla lista A-Z degli e-journal, che il Servizio bibliotecario pisano già produceva utilizzando le funzionalità di esportazione delle varie piattaforme e integrandovi conoscenze locali.[48]

Il *resolver*, in maniera trasparente per l'utente e prima di presentare il menu, consulta tale lista, e si comporta come segue:
– nei casi in cui ISSN e annate corrispondano (cioè i metadati della citazione vengano reperiti nella lista dei periodici elettronici) presenta l'"accesso diretto all'articolo" elettronico;
– se l'ISSN corrisponde ma l'anno della citazione non rientra nel *range* di abbonamento riportato nella lista, oppure se nella OpenURL manca il metadato dell'anno, il sistema segnala l'incongruenza ma propone ugualmente la homepage della rivista per controllo;
– se nella OpenURL manca il metadato ISSN, il sistema presenta la possibilità di ricerca per titolo nell'elenco A-Z dei periodici elettronici (nella versione su database online). (Figura 1)

La scelta di questo percorso è basata su due considerazioni: da un lato le dimensioni relativamente contenute della biblioteca digitale di riferimento rendono preferibile una certa ridondanza rispetto al rischio di perdere informazioni per errori o incompletezza dei dati (è questa, con tutta evidenza, una decisione strettamente condizionata dalla situazione locale); dall'altro, nella prospettiva di utilizzare lo strumento come una sorta di "reference indiretto", abbiamo ritenuto più "educativo" presentare sempre dall'utente l'intero percorso: obiettivo desiderabile, nel contesto dato, ci è apparso, non meno dell'accesso "con un solo clic", l'individuazione e il suggerimento di percorsi di ricerca corretti e completi.

Con "accesso diretto all'articolo" ci riferiamo al *tentativo di condurre all'articolo*, essendone la possibilità condizionata dalle specifiche tecniche della piattaforma dell'editore; nei casi in cui è tecnicamente impossibile arrivare proprio all'articolo, viene proposto l'accesso al livello (volume o fascicolo) più vicino possibile ad esso.

Per attivare questa funzionalità di *deep linking* (link a livello dell'articolo) è stata svolta una ricognizione delle possibilità offerte dalle diverse piattaforme, sulla base della documentazione messa a disposizione da ciascuna, o anche, quando necessario, con contatti diretti coi fornitori.

Abbiamo con soddisfazione riscontrato l'affermarsi di prassi standard da parte dei fornitori veramente interessati ad offrire il servizio di *deep linking*: infatti le sintassi di *link-to* incontrate rientrano quasi tutte in una delle seguenti tipologie:
– accesso tramite DOI;
– accesso tramite SICI, che risulta particolarmente elegante quando la URL può essere costruita modularmente coi metadati a disposizione e dunque funziona anche con metadati incompleti conducendo esattamente al livello che essi rappresentano (di testata, di volume, di fascicolo, di articolo);
– accesso tramite "*resolver* di ingresso", molto organico rispetto alla tecnologia in discussione.[49]

Per quanto riguarda l'accesso ai testi pieni di periodici *free-online*, si è ritenuto per ora sufficiente proporre la ricerca su DOAJ:[50] tale sito offre anche il servizio di estra-

[48] I dati scaricati dalle diverse piattaforme vengono passati in un database onde elaborarli e omogeneizzarli, poi dal relativo file vengono tratti vari prodotti (anche da mettere su web come accesso sintetico ai contenuti elettronici), tra cui un file di testo in ordine alfabetico (la lista A-Z appunto) che contiene un periodico per linea, e per ogni linea: titolo, ISSN, eISSN, URL della rivista, piattaforma, anno di inizio dei contenuti elettronici, anno di fine.

[49] Il generalizzarsi dell'OpenURL *resolver* d'ingresso è una via individuata pragmaticamente per far rientrare nell'ambito dello standard anche il colloquio lato *target* che a rigore ne è fuori; gli sviluppatori dello standard paiono recepire il suggerimento se il nono paragrafo di *The Key/Encoded-Value (KEV) format-implementation guidelines*, cit., si intitola (e promette): "Guidelines for OpenURL 'link-to' resolvers". Ciò che accade è che una risorsa *target* implementa un elementare *resolver* (per cui il *resolver* istituzionale si limita a passare la OpenURL che ha ricevuto, eventualmente arricchita), il quale gestisce come unico servizio l'accesso ai testi pieni del proprio *repository*, evitando così di pubblicizzare la propria specifica sintassi di accesso. Questo sistema ha evidenti vantaggi di semplicità e organicità, sebbene non abbia la medesima unicità di indirizzamento di una stringa costruita a mezzo identificatore e di fatto introduca un filtro ulteriore davanti alla procedura di accesso vero e proprio al database.

[50] Si tratta della Directory of Open Access Journals, repertorio internazionale di periodici online *peer-reviewed* a libero accesso: <http://www.doaj.org/doaj>.

zione della lista di ISSN e titoli da aggiungere alla lista A-Z degli abbonamenti dell'istituzione, ma data la dinamicità dei dati è parso preferibile mantenere la ricerca sulla directory, anche in questo caso scegliendo il percorso "più lungo ma più sicuro".[51]

Per quanto riguarda le monografie ad abbonamento *on-line*, numericamente minoritarie, non è stato necessario per ora prevedere percorsi di ricerca differenziati ma è stata ritenuta sufficiente la mediazione del catalogo generale di ateneo, su cui sono segnalate con relativo link.

Per il reperimento di monografie elettroniche *free* (in particolare dei testi classici messi a disposizione in varie forme dai diversi progetti di digitalizzazione) ci si affida alla ricerca tramite Google e Google Ricerca Libri.[52]

Due *target* di ambito particolare, presenti per tutti i tipi di documento, sono:

1) la possibilità di invio di una richiesta di ILL o DD alla propria biblioteca di riferimento;

2) la possibilità di invio di una richiesta di informazioni al servizio "Ask a librarian" dello SBA.

Il primo servizio è rivolto agli utenti istituzionali. Il sistema reindirizza:

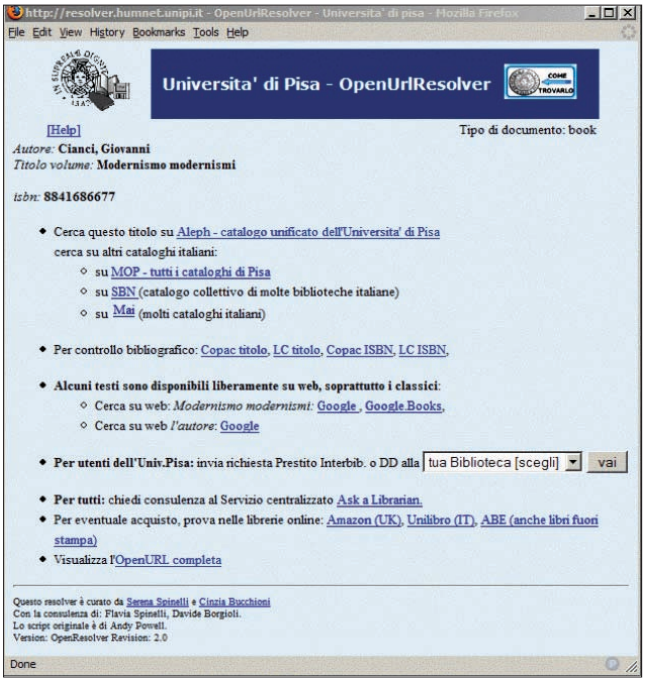
– per le biblioteche del circuito Nilde, alla *form*-utente di Nilde, precompilata;[53]

– per le altre biblioteche, presenta una *form* precompilata che accetta richieste corredate di indirizzo e-mail e matricola (o codice di iscrizione al prestito) e le invia per posta elettronica alla biblioteca scelta, demandandole i controlli del caso.[54]

Il secondo servizio è rivolto agli utenti esterni e dunque privi di una "biblioteca di ateneo di riferimento", *walk-in users* o utenti che abbiano trovato su web la *form* OpenURL *generator*[55] (il *resolver* non è di per sé ad accesso riservato).

Abbiamo infine impostato nel *resolver* la possibilità di differenziare i menu a seconda dell'indirizzo IP di provenienza della chiamata (dell'utente): questa funzionalità è testata (e impostata) a livello implementativo, ma per adesso non è attivata, in quanto manca un'analisi della sua effettiva utilità a fronte della complicazione della gestione di menu differenziati. L'ipotesi è che potrebbe essere usata per differenziare i servizi messi a disposizione dei bibliotecari rispetto a quelli proposti agli utenti finali (aggiungervi link a biblioteche straniere per ILL e DD; a siti per acquisti ecc.); o anche per differenziare i menu a seconda dei settori disciplinari.[56] (Figure 2 e 3)

Fig. 2 - Esempio di menu di servizi per una monografia, con riga aggiuntiva di servizi "per controllo bibliografico"



[51] Invece i *free-journals* di interesse, ma non inclusi in DOAJ, vengono quando reperiti aggiunti alla lista A-Z. Aggiungiamo l'ovvia osservazione che l'attivazione di un servizio di questo tipo a livello di ateneo contribuisce a far emergere la criticità di una gestione frazionata tra livello centrale di Sistema bibliotecario e singole biblioteche; diventa infatti molto più difficile valorizzare gli strumenti "della biblioteca", off-line o ad abbonamento riservato per ragioni di risparmio o anche di semplice tradizione. Soluzioni di questo tipo pongono con forza l'esigenza di maggior sistematicità, o almeno coordinamento e comunicazione.

[52] Come si nota nelle schermate riportate, si è cercato di sottolineare nel menu la possibilità di recupero dei testi pieni tramite motori di ricerca generici. Solo in grazia di questa possibilità è stata presa in considerazione l'opportunità di attivare come *source* il catalogo di ateneo, peraltro tipicamente un *target*. Sulla base delle stesse considerazioni è parso utile inserire come *source* anche il MOP (MetaOpac Pisano), sistema a gestione locale curato da Giuseppe Romano, nel quale è stata appositamente implementata la conformità allo standard con emissione di OpenURL.

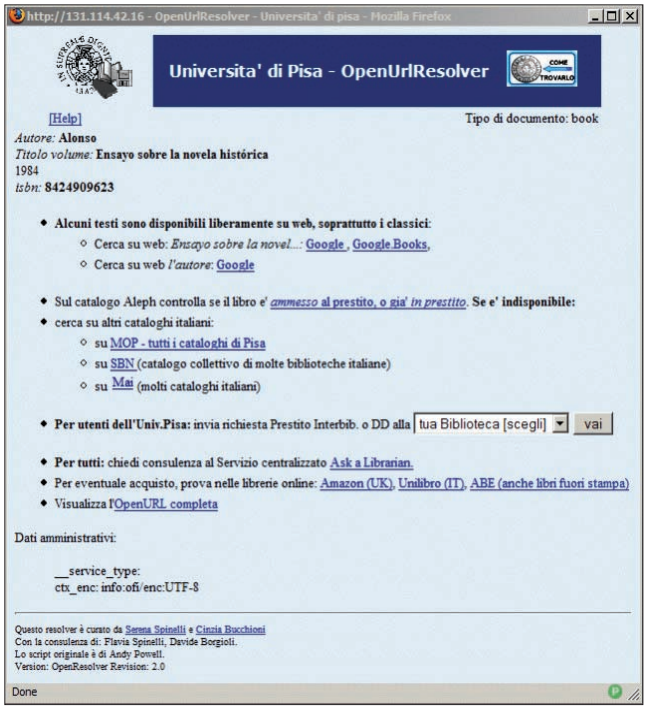
[53] Il sistema Nilde consente agli utenti registrati di inoltrare le richiesta tramite una *form* sulla sua piattaforma, ma è anche in grado di ricevere un'OpenURL, avendo implementato a sua volta un *resolver*; al momento della chiamata chiede all'utente di autenticarsi sul sistema.

[54] Questo perché nelle biblioteche che non partecipano a Nilde non sono in uso sistemi specifici di ricezione delle richieste ILL/DD, tranne in un caso in cui è in uso un sistema casalingo rispetto al quale è in corso uno studio per l'integrazione.

[55] Cfr. nota 43.

[56] Il presente servizio a gestione centralizzata ha concorso a far emergere le differenze di percorsi di ricerca, o almeno di enfasi, nei vari ambiti disciplinari; certo la distinzione per numero IP non è del tutto congrua allo scopo di distinguere gli ambiti disciplinari, identificando semplicemente la sede fisica dell'università da cui si collega l'utente richiedente il servizio. Se una seria analisi costi/benefici sull'opportunità di differenziare l'offerta di risorse e servizi per disciplina risultasse opportuna, l'implementazione dovrebbe fare piuttosto uso di dispositivi di autenticazione dell'utente, eventualmente introducendo nella OpenURL l'entità che lo identifica (*requester*, cfr. nota 60). Parrebbe comunque il caso di fermarsi ad una categorizzazione a maglie molto larghe; e se è vero che una specializzazione disciplinare del menu del *resolver* permetterebbe di valorizzare risorse specifiche, è anche vero che in assenza di sistemi condivisi e interoperabili per l'indicizzazione delle risorse, il mantenimento di una *knowledge base* locale che gestisca anche questo aspetto si prospetta molto oneroso.

Fig. 3 - Esempio di menu di servizi per una monografia, la cui OpenURL è inviata dal catalogo di ateneo; si noti il menu configurato per la ricerca su web e per il controllo dello stato di prestito sul sistema locale



2.6. La knowledge base

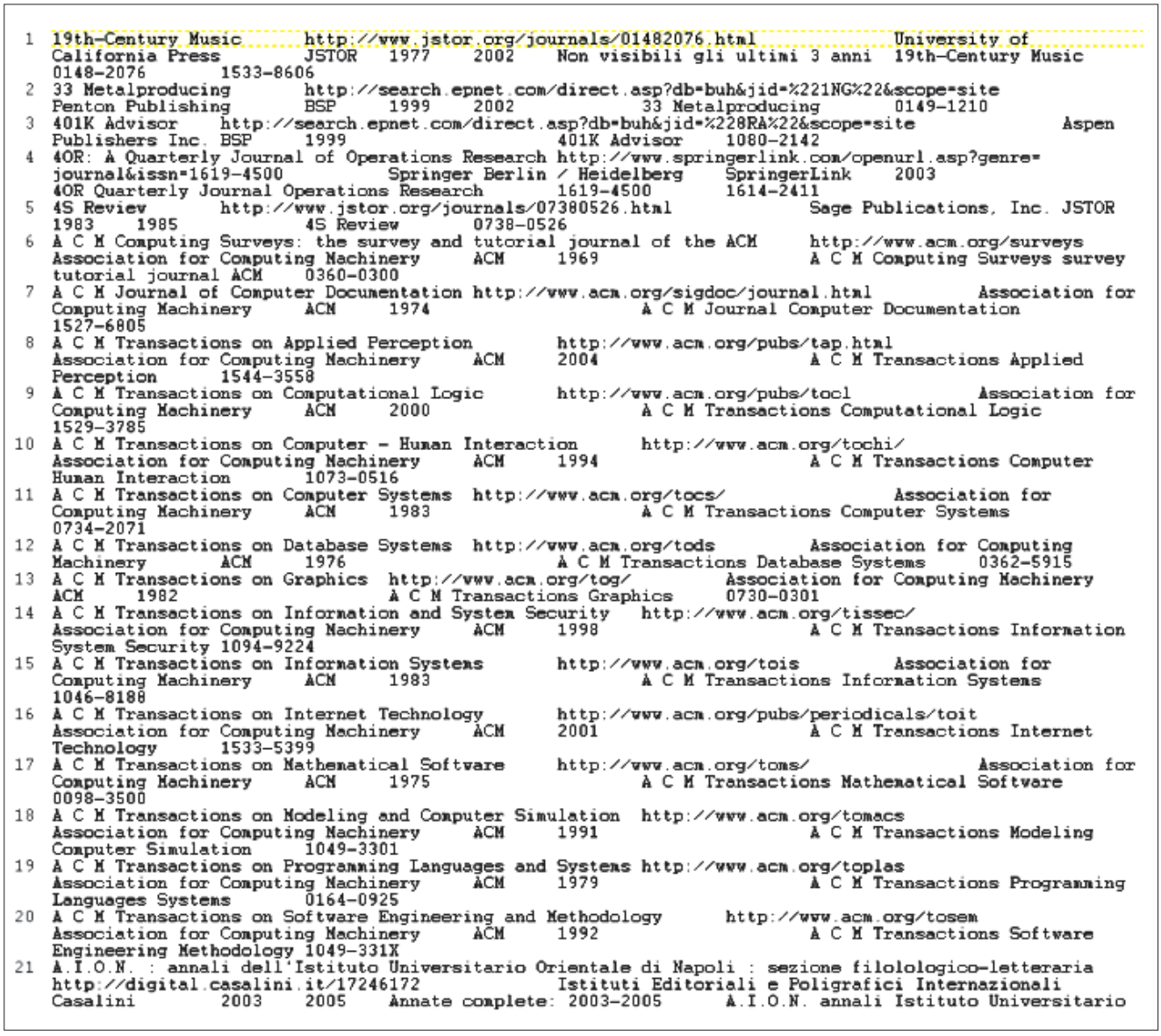
Uno degli elementi classicamente citati a proposito degli OpenURL *resolver* è la *knowledge base*. Essa costituisce di fatto la componente di maggior valore di un *resolver*, ma si tratta di un concetto di una qualche ambiguità: in generale la si può intendere come "sistema di informazioni a disposizione di un *resolver* per il suo funzionamento".
Nel caso in esame, la *knowledge base* può considerarsi costituita essenzialmente dai seguenti elementi:
1) la lista A-Z dei periodici elettronici (figura 4);
2) le procedure implementative per ottimizzare i rapporti con le risorse *source* (*source-parser*) e per accedere ai *target* (*target-parser*);
3) i meccanismi di arricchimento dei metadati;
4) le logiche procedurali di selezione dei servizi e costruzione dei menu.
In questo senso, dunque, essa non ci pare riferibile ai soli elementi repertoriali ben identificabili all'interno del sistema (per intenderci, il punto 1), ma necessariamente inclusiva di logiche e procedure più o meno integrate nello *script*. È evidente che un sistema è tanto più esportabile quanto più riesce a isolare i parametri variabili e a gestirli al di fuori del codice di programmazione: i sistemi commerciali, a massima generalizzazione, si caratterizzano per l'impegno in questa direzione (cioè di massima separazione della *knowledge base* dal codice di programmazione, e gestione di essa tramite opzioni di configurazione).
Il sistema qui presentato non ha evidentemente specifiche ambizioni in questo senso; ripercorrendo gli elementi sopra elencati, la lista A-Z è un file di testo a sé stante, e va in ogni caso prodotta localmente. Sul versante *source-parser* e *target-parser* sono stati affrontati solo i casi di interesse contestuale senza pretese di esaustività, anche se questo ha di fatto significato prendere in carico le più diffuse risorse commerciali e i principali cataloghi italiani: su questo versante c'è l'intenzione di procedere a una maggior modularizzazione e parametrizzazione per facilitare l'aggiunta di nuovi elementi.
Con riferimento ai punti 3 e 4, ci pare che essi innegabilmente incorporino aspetti e scelte relative al contesto, tali però da risultare necessariamente inestricabili dalla logica del programma e difficilmente traducibili in opzioni di configurazione pre-stabilite all'interno di un software chiuso.
Se dunque è ovvio che l'ingegnerizzazione del software, a fronte dei vantaggi di pre-confezione, comporta una qualche limitazione delle possibilità di personalizzazione, ci pare che lo strumento in oggetto abbia caratteristiche tali (dipendenza dal contesto applicativo, dipendenza dall'incostanza dei servizi di rete, basilare semplicità implementativa) da risultare particolarmente adatto alla filosofia "artigianale" del software *open-source*.[57]

3. Dall'applicazione alla visione: prospettive e sviluppi

In chiusura ci piace fare un breve accenno al dibattito e alle proposte più recenti che negli ambienti delle tecnologie della digital library riguardano in qualche modo la OpenURL.
Perché se è vero che dal punto di vista delle applicazioni le uniche a regime paiono ancora i *resolver*, sperimentazioni piuttosto varie sono in campo da anni: si ricordi l'interessante rassegna di utilizzi emergenti presentata da Ann Apps già al Convegno "ELPUB2005 The Ninth ICCC International conference on electronic publishing – From author to reader", tenutosi presso la Katholieke Universiteit di Leuven, nei giorni 8-10 giugno 2005.[58]
Si possono forse individuare due principali linee di sviluppo. La prima è riconducibile all'interesse a mettere a frutto le ampie potenzialità dello standard, dunque ad utilizzarne la valenza descrittiva;[59] ad estenderne l'applicazione a nuovi domini, introducendo nel registro nuovi set di metadati; ad utilizzare effettivamente le entità contestuali diverse dal-

[57] Dato il codice sorgente, possibilmente documentato, si tratta non di configurarlo secondo parametri prestabiliti, ma di integrarlo e svilupparlo, intervenendo direttamente nello *script*, ovviamente da parte di un esperto informatico.
[58] Ann Apps – Ross MacIntyre, *Emerging uses for the OpenURL framework*, <http://epub.mimas.ac.uk/papers/elpub2005/appsmac-elpub2005_full.html>.
[59] Si pensi ad esempio all'inserimento delle informazioni citazionali all'interno di una descrizione in formato Dublin Core. Cfr. *Guidelines for encoding bibliographic citation information in Dublin Core metadata*, <http://dublincore.org/documents/dc-citation-guidelines/>.

Fig. 4 - Porzione di lista A-Z



```
1  19th-Century Music      http://www.jstor.org/journals/01482076.html              University of
   California Press        JSTOR    1977    2002    Non visibili gli ultimi 3 anni  19th-Century Music
   0148-2076       1533-8606
2  33 Metalproducing       http://search.epnet.com/direct.asp?db=buh&jid=%221NG%22&scope=site
   Penton Publishing       BSP      1999    2002            33 Metalproducing       0149-1210
3  401K Advisor    http://search.epnet.com/direct.asp?db=buh&jid=%228RA%22&scope=site               Aspen
   Publishers Inc. BSP     1999                    401K Advisor    1080-2142
4  4OR: A Quarterly Journal of Operations Research http://www.springerlink.com/openurl.asp?genre=
   journal&issn=1619-4500          Springer Berlin / Heidelberg    SpringerLink    2003
   4OR Quarterly Journal Operations Research       1619-4500       1614-2411
5  4S Review       http://www.jstor.org/journals/07380526.html             Sage Publications, Inc. JSTOR
   1983    1985            4S Review       0738-0526
6  A C M Computing Surveys: the survey and tutorial journal of the ACM     http://www.acm.org/surveys
   Association for Computing Machinery     ACM     1969                    A C M Computing Surveys survey
   tutorial journal ACM    0360-0300
7  A C M Journal of Computer Documentation http://www.acm.org/sigdoc/journal.html          Association for
   Computing Machinery     ACM     1974                    A C M Journal Computer Documentation
   1527-6805
8  A C M Transactions on Applied Perception        http://www.acm.org/pubs/tap.html
   Association for Computing Machinery     ACM     2004                    A C M Transactions Applied
   Perception      1544-3558
9  A C M Transactions on Computational Logic       http://www.acm.org/pubs/tocl            Association for
   Computing Machinery     ACM     2000                    A C M Transactions Computational Logic
   1529-3785
10 A C M Transactions on Computer - Human Interaction      http://www.acm.org/tochi/
   Association for Computing Machinery     ACM     1994                    A C M Transactions Computer
   Human Interaction       1073-0516
11 A C M Transactions on Computer Systems  http://www.acm.org/tocs/                Association for
   Computing Machinery     ACM     1983                    A C M Transactions Computer Systems
   0734-2071
12 A C M Transactions on Database Systems  http://www.acm.org/tods         Association for Computing
   Machinery       ACM     1976                    A C M Transactions Database Systems     0362-5915
13 A C M Transactions on Graphics  http://www.acm.org/tog/         Association for Computing Machinery
   ACM     1982            A C M Transactions Graphics     0730-0301
14 A C M Transactions on Information and System Security   http://www.acm.org/tissec/
   Association for Computing Machinery     ACM     1998                    A C M Transactions Information
   System Security 1094-9224
15 A C M Transactions on Information Systems       http://www.acm.org/tois         Association for
   Computing Machinery     ACM     1983                    A C M Transactions Information Systems
   1046-8188
16 A C M Transactions on Internet Technology       http://www.acm.org/pubs/periodicals/toit
   Association for Computing Machinery     ACM     2001                    A C M Transactions Internet
   Technology      1533-5399
17 A C M Transactions on Mathematical Software     http://www.acm.org/toms/                Association for
   Computing Machinery     ACM     1975                    A C M Transactions Mathematical Software
   0098-3500
18 A C M Transactions on Modeling and Computer Simulation  http://www.acm.org/tomacs
   Association for Computing Machinery     ACM     1991                    A C M Transactions Modeling
   Computer Simulation     1049-3301
19 A C M Transactions on Programming Languages and Systems http://www.acm.org/toplas
   Association for Computing Machinery     ACM     1979                    A C M Transactions Programming
   Languages Systems       0164-0925
20 A C M Transactions on Software Engineering and Methodology      http://www.acm.org/tosem
   Association for Computing Machinery     ACM     1992                    A C M Transactions Software
   Engineering Methodology 1049-331X
21 A.I.O.N. : annali dell'Istituto Universitario Orientale di Napoli : sezione filolologico-letteraria
   http://digital.casalini.it/17246172             Istituti Editoriali e Poligrafici Internazionali
   Casalini        2003    2005    Annate complete: 2003-2005      A.I.O.N. annali Istituto Universitario
```

l'oggetto citato (*referent*).[60] Su questo versante si collocano alcuni dei progetti più recenti dello stesso Van de Sompel che impiegano la OpenURL per il colloquio *server-to-server* ed esplorano utilizzi congiunti di OAI-PMH[61] e OpenURL.[62]

Una seconda linea di sviluppo è quella che si può intravedere nelle sperimentazioni in corso nell'ambito di alcuni gruppi di ricerca più lontani dagli originari sviluppatori, in cui si rilevano piuttosto tentativi di lettura (e forse di forzatura) dello standard alla luce di filosofie di-

[60] Si ricorda che le entità che possono essere descritte nel ContextObject, secondo lo standard, sono: un oggetto citato e vari elementi che definiscono il contesto di rete in cui avviene la citazione. Le entità definite sono dunque sei:
– *Referent*: oggetto (pubblicazione) citata;
– *ReferringEntity*: pubblicazione o record in cui si trova la citazione;
– *Referrer*: piattaforma o servizio in cui si trova la citazione;
– *Requester*: utente finale;
– *ServiceType*: tipo di servizio (eventualmente specificato: full text o posseduto a catalogo);
– *Resolver*: OpenURL resolver.

[61] Open Archives Initiative Protocol for Metadata Harvesting, <http://www.openarchives.org/OAI/openarchivesprotocol.html>.

[62] Herbert Van de Sompel et al., *An interoperable fabric for scholarly value chains*, "D-Lib Magazine", 12 (2006), 10, <http://www.dlib.org/dlib/october06/vandesompel/10vandesompel.html>.

verse, orientate allo stile del web 2.0. Si può in qualche modo considerare un primo passo in questa direzione anche la proposta dei già citati COinS (cfr. nota 19): l'idea di inserire i metadati della OpenURL nel codice html, invisibili (*latent*) fintanto che non intervenga un dispositivo locale di *auto-discovery*,[63] cioè un'estensione del browser opportunamente settata dall'utente col *resolver* di propria scelta, supera l'idea dell'uso istituzionale per andare verso una concezione del web come "spazio" personalizzabile. Ancora oltre in questa direzione va il dibattito in corso su alcuni blog e liste di discussione, per esempio le suggestioni lanciate da Dan Chudnov e variamente raccolte:[64] partendo dall'osservazione della diffusione dei "*resolver* d'ingresso"[65] (cioè interfacce a singoli servizi attivate da una OpenURL: per esempio RefWork, Nilde e molti fornitori di contenuti si interfacciano ai *resolver* facendosi passare direttamente la OpenURL e su tale base forniscono il proprio servizio) si mette in dubbio la sostenibilità computazionale a largo raggio della concezione *context sensitive* e si propone la OpenURL come dispositivo di accesso ai singoli servizi, in uno scenario in cui sia l'utente finale a poterli – in qualche modo da definirsi – selezionare in funzione del proprio personale contesto. Una visione insomma fuori da ogni "gestione del bibliotecario", evidentemente orientata a un *target* di utenti molto più *net generation* di quello cui attualmente rivolgiamo i nostri servizi, che può essere suggestiva anche se forse troppo visionaria, e certo per ora puramente teorica.

Abstract

The paper includes three parts. The first part offers a review of the debate on reference linking started in the 90ties, of the context, the discussions and the achievements of it, from the emerging of the "appropriate copy" problem, through the proposal of permanent identifiers, until the rising of the OpenURL solution, which appears rather close to the library tradition since proposes an indirect linking based on a standard description and aimed at the delivery of extended services to final users.

The second part of the paper focuses on a case study, the home grown OpenURL resolver at the University of Pisa: a linking tool under library control seemed a desirable addition to the Pisa library system, where no unique catalogue access to the different print and digital resources was available; it was self developed since linking tools appear to be more effective the more they deal with the local situation (i.e. the more they support customization), and in this case orientation toward the local situation has meant orientation toward the hybrid library, i.e. attention to opacs, traditional services, and not only to identifiers but also to bibliographic metadata. The application has been developed from the Perl script of Andy Powell known as UKOLN resolver, with relevant improvements: compliance to OpenURL 1.0, profile SAP1, interoperability with the local A-Z list of e-journals, inclusion of sources and targets actually subscribed to or used at the University of Pisa. It is not intended for generalization but for availability to the developer community under the GPL licence.

The third part shortly traces a review of the newest applications of the OpenURL standard in the context of web 2.0.

[63] Daniel Chudnov et al., *Opening up OpenURLs with autodiscovery*, "Ariadne", 43 (aprile 2005), <http://www.ariadne.ac.uk/issue43/chudnov/>.

[64] Si veda, sul blog di Jon Udell, il post *A conversation with Dan Chudnov about OpenURL, context-sensitive linking, and digital archiving*, 16 febbraio 2007, <http://blog.jonudell.net/2007/02/16/a-conversation-with-dan-chudnov-about-openurl-context-sensitive-linking-and-digital-archiving/>; e le successive considerazioni, sul blog "Bibliographic Wilderness", del post di Jonathan Rochkind, *"Link resolver" understood as "OpenURL service provider"*, 6 marzo 2007, <http://bibwild.wordpress.com/2007/03/06/link-resolver-understood-as-openurl-service-provider/>.

[65] Cfr. nota 49.

Appendice 1

OpenURL Resolver Toolkit

Per consentire il test del *resolver* abbiamo predisposto alcuni strumenti:
– una pagina per impostare alcuni *resolver*, fra cui in particolare quello illustrato in questo articolo, su diverse risorse *source*, tramite "cookiepusher": <http://resolver.humnet.unipi.it/>;
– una pagina attraverso la quale testare il funzionamento dei *resolver* di cui sopra con diverse tipologie di OpenURL: <http://resolver.humnet.unipi.it/testcook3.htm>;
– una pagina da cui scaricare e provare alcuni *bookmarklet*, tra cui quelli personalizzati per l'OpenURL Resolver: <http://resolver.humnet.unipi.it/testcoins.htm>.
Per agevolare i futuri sviluppi del *resolver*, la discussione e la partecipazione degli eventuali interessati abbiamo predisposto un blog, disponibile su <openurlresolver.blogspot.com>.
Il *resolver* verrà distribuito con licenza GNU GPL. Il testo della licenza è disponibile su <http://www.gnu.org/licenses/gpl.txt>. La traduzione italiana della licenza è disponibile su <http://www.softwarelibero.it/gnudoc/gpl.it.txt>.

Appendice 2

COinS - ContextObject in Span

... ovvero: una proposta di standard per rendere *OpenUrl compliant* (capace di emettere una OpenUrl) anche una pagina web statica:

1) inserire la descrizione OpenURL nel codice html, all'interno dell'elemento , che non viene normalmente visualizzato dai browser, con una sintassi del tipo:

```
<span class="Z3988" title="ctx_ver=Z39.88-
2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&rft_id=doi:10.1108/
02640470410561893&rft.title=The Electronic
Library&rft.atitle=Weblogs%20-%20are you serious?"></span>
```

2) utilizzare apposite estensioni o plug-in del browser locale per trasformare tale descrizione in link al resolver:
a) con il browser FireFox, si può installare la *extension* OpenReferrer, la quale permette di settare, così come la homepage, anche il *resolver* di riferimento;
b) con il browser Internet Explorer, si può installare un *bookmarklet* (cioè un bottoncino sulla barra che nasconde un programmino), cliccando il quale sulle pagine che contengono COinS si ottiene la visualizzazione dei link al *resolver*.

Sulla proposta COinS: <http://ocoins.info/>.

Per scaricare l'*extension* per Firefox: <http://www.openly.com/openurlref/>.

Per testare e installare il *bookmarklet* dell'OpenUrlResolver: <http://resolver.humnet.unipi.it/ testcoins.htm>.

Per provare l'*extension* o il *bookmarklet* su altre pagine che implementano COinS: <http://ocoins.info/#id3205609424>.

Appendice 3

La OpenURL

Nella sua struttura più semplice e più diffusa è un particolare tipo di URL (formato GET) e ha dunque la struttura definita dal W3C per tutte le URI: protocollo://server.di.rete/dirEfile?query.
Ad esempio: <http://www.google.it/search?hl=it&q=prova>.
La *query*, che è opzionale, è ignorata dal web e consegnata per l'interpretazione al server/programma chiamato, dopo essere stata sottoposta alla codifica (*url-encoding*) dei caratteri per i quali ciò è necessario (per es. %20 al posto di *spazio*).

Esempio di OpenURL 0.1

<http://example.org/myResolver?genre=article&aulast=Santone&aufirst=Antonella&atitle=
Automatic%20verification%20of%20concurrent%20systems%20using%20a%20formula-
based%20compositional%20approach&title=Acta%20Informatica&date=2002&
volume=38&issue=8&spage=531&issn=0001-5903>

La stringa in rosso è l'indirizzo di rete del resolver locale; segue (dopo ?) la *query* codificata secondo lo standard OpenURL e dunque contenente la descrizione di una pubblicazione espressa in un elenco di parametri (divisi da "&") che sono metadati bibliografici, nella forma: nome=valore.

Esempi di OpenURL 1.0 conformi al profilo SAP1

Si differenziano dalla precedente perché includono varie informazioni amministrative e autodichiarano i formati utilizzati, selezionati tra quelli elencati nel registro.

Esempio 1

<http://example.org/myResolver?**url_ver=Z39.88-2004&url_ctx_fmt=info:ofi/fmt: kev:
mtx:ctx&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&ctx_ver=Z39.88-2004 &ctx_enc=
info:ofi/enc: ISO-8859-1**&rft.volume=12&rft.issn=0306-
1078&rft.genre=article&rft.issue=3&rft.title=Early%20music&rft.aufirst=Peter&rft.atit
le=Violin%20Fingering%20in%20the%2018th%20Century&rft.aulast=Walls&rft.spage=299**&rfr_
id=info:sid/copac.ac.uk>**

Questa particolare istanza:
– utilizza per il ContextObject il formato a coppie nome=valore, identificato nel registro come: *info:ofi/fmt:kev:mtx:ctx*;
– utilizza per la descrizione del ***referent*** i metadati definiti per i periodici (identificati nel registro come: *info:ofi/fmt:kev:mtx:journal*);
– utilizza per la descrizione dell'entità (facoltativa) ***referrer*** un identificatore (***rfr_id***);
– è codificata in URL come "*inline OpenURL*" (senza doppia codifica, come la precedente).

Esempio 2

<http://example.org/myResolver?url_ver=**Z39.88-2004**&url_tim=**2004-04-
23&url_ctx_fmt=info%3Aofi%2Ffmt%3Akev%3Amtx%3Actx&url_ctx_val=**ctx_ver%3DZ39.88-
2004%26ctx_enc%3Dinfo%253Aofi%252Fenc%253AUTF-8%26ctx_id%3D
345871%26ctx_tim%3D2002-03-20T08%253A55%253A12Z%26rft_val_fmt%3D
info%253Aofi%252Ffmt%253Akev%253Amtx%253Ajournal%26rft.atitle%3DIsolation%2520of%2520
a%2520common%2520receptor%2520for%2520coxsackie%2520B%26rft.jtitle%3DScience%26rft.au
last%3DBergelson%26rft.auinit%3DJ%26rft.date%3D1997%26rft.volume%3D275%26rft.spage%3D
1320%26rft.epage%3D1323**%26rfe_val_fmt%3Dinfo%253Aofi%252Ffmt%253Akev%253Amtx%253Ajour
nal%26rfe.atitle%3Dp27-p16%2520
Chimera%253A%2520A%2520Superior%2520Antiproliferative%26rfe.jtitle%3D
Molecular%2520Therapy%26rfe.aulast%3DMcArthur%26rfe.aufirst%3DJames
%26rfe.date%3D2001%26rfe.volume%3D3%26rfe.issue%3D1%26rfe.spage%3D8%26rfe.epage%3D13%
26req_ref_fmt%3Dhttp%253A%252f%252Flib.caltech.edu%252Ffmt%252Fldap-
mtx.html%26req_ref%3Dhttp%253A%252F%252Fldap.caltech.edu%252Fjaned%252Frecord.txt>

Questa particolare istanza:
– utilizza per il contextObject il formato a coppie *nome=valore*, identificato nel registro come: *info:ofi/fmt:kev:mtx:ctx* (come sopra);
– utilizza per la descrizione del ***referent*** i metadati definiti per i periodici (identificati nel registro come: *info:ofi/fmt:kev:mtx:journal*);
– utilizza per la descrizione dell'entità (facoltativa) ***referring entity*** i medesimi metadati;
– la descrizione dell'entità (facoltativa) ***requester*** è *by-reference* (o per citazione) e dunque viene riportata non la descrizione ma la URI dove è reperibile: <*http://ldap.caltech.edu/janed/record.txt*>; il formato di descrizione è a sua volta definito alla URI: <http://lib.caltech.edu/fmt/ldap-mtx.html>.
– è codificata in URL come "*by-value OpenURL*" (l'intero ContextObject è espresso come valore di **url_ctx_val=** e dunque sottoposto a doppia codifica).