

Catalogando, catalogando ...metacatalogando

*Come può cambiare una funzione
tradizionale della biblioteca*

Fino ad oggi le biblioteche hanno descritto tutti i libri ed i periodici presenti nella loro collezione attraverso il catalogo. In genere non sono stati catalogati gli articoli dei periodici, compito che è stato demandato alle bibliografie ed alle banche dati. Perciò i cataloghi, le bibliografie e le banche dati sono i tre strumenti con cui le biblioteche consentono l'accesso alle fonti informative. Il catalogo in particolare descrive ciò che la biblioteca ha nella sua collezione, mentre bibliografie e banche dati sono indipendenti dalla raccolta della biblioteca, e descrivono tutto ciò che è pubblicato in un'area geografica o su uno specifico argomento. Una volta identificato il documento nelle bibliografie o nelle banche dati, l'utente deve consultare il catalogo della biblioteca o un catalogo unico per sapere dov'è conservato il documento mentre per accedere ai documenti l'utente deve andare necessariamente in biblioteca.

Un servizio che non c'è

Con i documenti elettronici in rete la situazione è molto più complessa. La rete consente l'accesso ad una moltitudine di risorse che non sono in biblioteca ed a cui l'utente accede direttamente. Gli utenti ora non dipendono più dalle biblioteche per ottenere i documenti di cui hanno bisogno. Di conseguenza le biblioteche non possono limitare la loro funzione di intermediazione alle fonti raccolte localmente ma,

se vogliono aiutare l'utente ad identificare e ad accedere anche a tutte le altre fonti, devono estendere questa funzione alle risorse in rete. In altre parole, questo è il vero significato degli slogan ora di moda "accesso e non possesso" oppure "just in case" e non "just in time".

C'è un'esigenza reale: l'utente è come in una giungla dove c'è di tutto, fonti di qualità e fonti con informazione di nessun valore, e deve essere aiutato a non perdere il suo tempo, trovando subito e solo ciò che ha valore per lui. Il servizio che manca è questo: un unico punto nella rete dove trovare tutto ciò di cui si ha bisogno con un'unica ricerca. Le biblioteche possono, se vogliono, fornire questo tipo di aiuto, ad esempio catalogando risorse elettroniche che non sono fisicamente presenti in biblioteca ma sono accessibili in rete.

Le risorse elettroniche in rete comprendono tre tipi di documenti:

— articoli in linea. Ad esempio Elsevier, Academic Press, Kluwer forniscono il contenuto delle riviste in formato elettronico. Il testo pieno dell'articolo è accompagnato da una parte del documento che contiene "metainformazione" sul testo e cioè l'intestazione e l'indice dei contenuti (in formato Sgml così da poter essere riversato in qualsiasi altro formato). La meta-informazione po-

trebbe essere usata per la produzione di citazioni bibliografiche nella banca dati del catalogo e queste citazioni potrebbero fare riferimento alla localizzazione del documento nel server accessibile da Internet; — documenti elettronici, in genere letteratura grigia: i rapporti di ricerca, spesso disponibili sia su carta che in versione elettronica, potrebbero essere descritti sia nelle banche dati che nel catalogo, con la Url (cioè la localizzazione in un sito Web) aggiunta in un campo apposito. L'utente dovrebbe però sapere quale programma è necessario per visualizzare il documento (ad esempio un browser di Web o un viewer per Pdf a seconda che il documento sia in formato Html o in un formato Pdf);

— i siti di Internet: alcuni dei siti accessibili in Internet hanno fonti bibliografiche di grande interesse per le biblioteche oppure funzionano come gateway ad una molteplicità di risorse in rete; in questo caso potrebbe essere creato un collegamento ipertestuale dall'opac al server Internet.

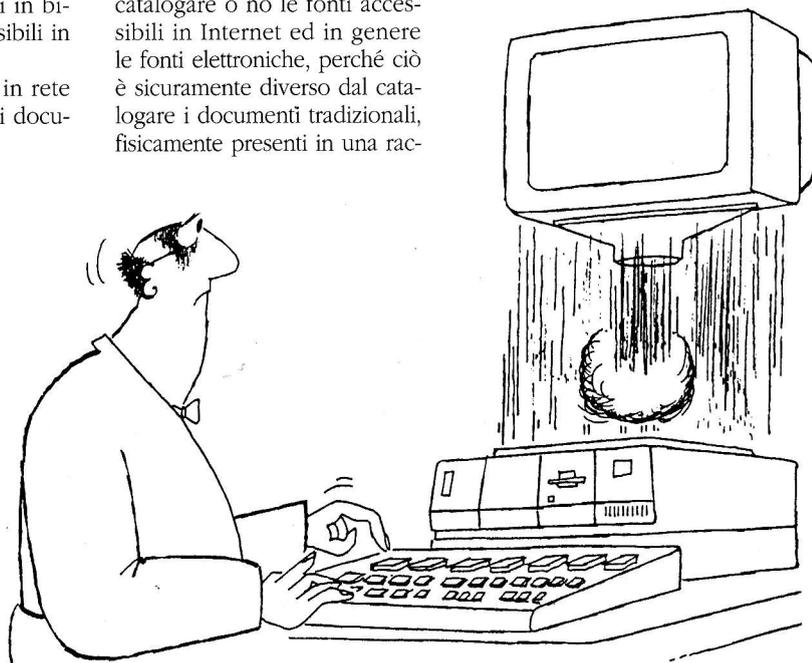
Molte sono attualmente le discussioni tra i bibliotecari se catalogare o no le fonti accessibili in Internet ed in genere le fonti elettroniche, perché ciò è sicuramente diverso dal catalogare i documenti tradizionali, fisicamente presenti in una rac-

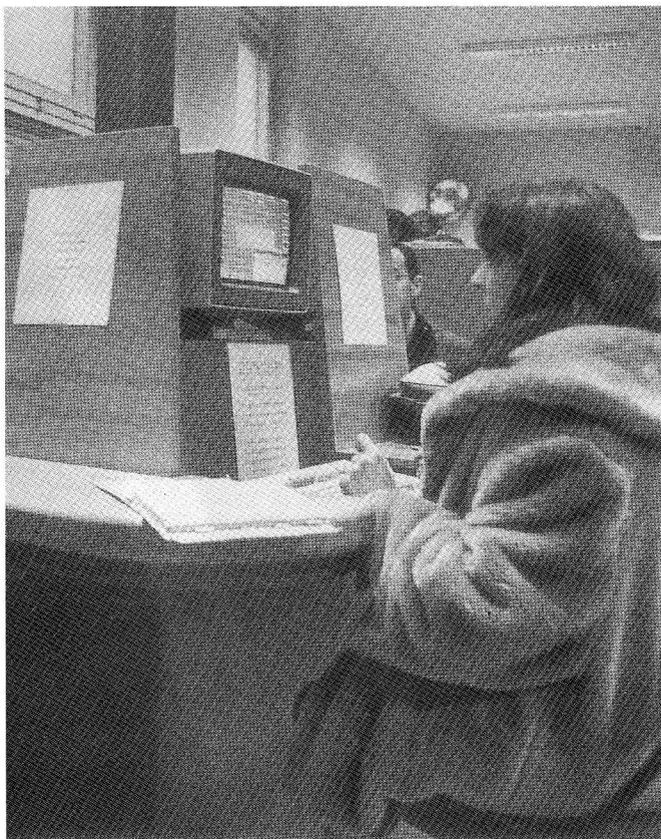
colta bibliotecaria. Difficile dare risposte certe in queste discussioni se non si ha chiarezza di come estendere le tradizionali funzioni di acquisizione e catalogazione alle fonti in rete. Soprattutto è difficile immaginare come queste funzioni devono cambiare, perché il peso della tradizione porta a forzare il nuovo nel vecchio, salvo poi accorgersi che ci va troppo stretto. Ma è stato sempre così. Quando si ha uno strumento nuovo, come è adesso Internet, cerchiamo sempre di fargli fare quello che si faceva prima, forse meglio, senza quello strumento.

Catalogare... o metacatalogare?

Semplificando molto, si può dire che ci sono quattro modi in cui la biblioteca può aiutare l'utente ad identificare ed accedere alle fonti informative di rete:

1) la biblioteca rende disponibile l'opac in Internet. Sempre più spesso gli opac includono la descrizione di documenti elettronici sia che questi siano posseduti dalla biblioteca sia





che siano memorizzati in server che la biblioteca controlla. Le regole di catalogazione sono state adattate a consentire la descrizione della localizzazione per consentire l'accesso al documento (Url). Tuttavia molti cominciano a dubitare della convenienza di tale pratica di catalogazione per la dinamicità (ossia l'estrema variabilità del contenuto insieme all'instabilità della disponibilità nel sito) dei documenti in Internet e l'immensità delle fonti: non è possibile mantenere a costi accettabili un catalogo aggiornato;

2) la biblioteca mette a disposizione una completa selezione dei molti elenchi esistenti di siti Gopher, Ftp, Www, ecc. sia memorizzando gli elenchi nell'opac (ma in questo caso non è possibile cliccare e navigare nel sito Internet) sia predisponendo dei collegamenti automatici (link) alla loro localizzazione nella rete, da tenere

sempre aggiornati. Molti sono gli sforzi per arrivare ad un modo di citazione uniforme delle risorse in Internet;

3) mantenere un sito Web, coordinato dalla biblioteca, che funziona come punto di accesso a Internet, organizzando nel modo più efficace per l'utenza (per soggetto, per tipo di documento, per area geografica, ecc.) sia le fonti reperibili localmente che le esterne (individuate da collegamenti (link) al sito Internet);

4) in aggiunta ad estensioni dell'opac per i documenti elettronici localizzati nella biblioteca o in siti esterni, la biblioteca organizza l'accesso ai vari motori di ricerca che oggi esistono su Internet e che indicizzano automaticamente la gran parte di ciò che esiste, consentendo di fare la ricerca per parole chiave ed altri identificatori come la classificazione Cdu (Www Virtual Library, Aliweb, Yahoo, ecc.); in questo caso

l'opac deve essere interrogabile con un browser standard di Web come interfaccia dell'utente e consentire sia la ricerca del catalogo sia tutte le funzionalità previste da Internet (telnet, Ftp, posta elettronica); una possibilità più avanzata è quella di avere come interfaccia un client Z39.50.¹

Le varie possibilità sono state elencate in ordine di difficoltà organizzative e tecnologiche di realizzazione e corrispondono all'evoluzione che hanno avuto i primi progetti di catalogazione delle fonti in rete. La scelta tra i vari metodi dipende a questo punto dalle risorse che si hanno.

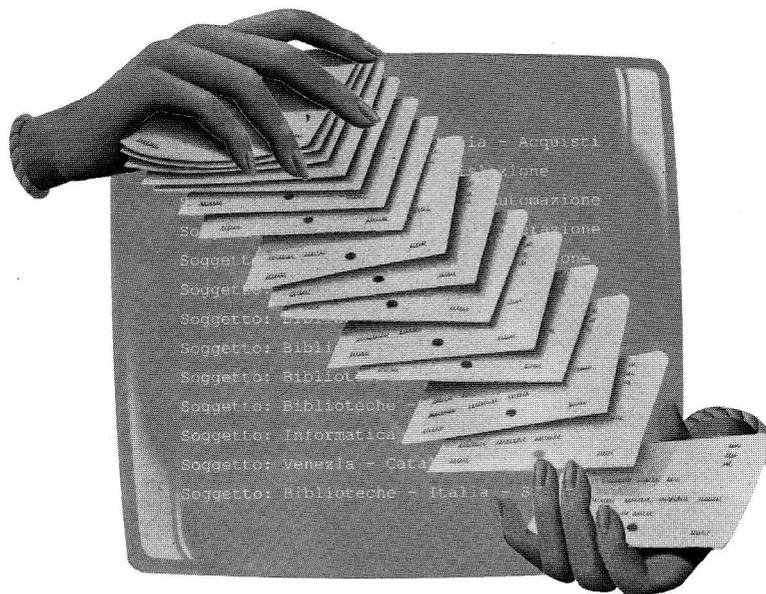
È da tener ben presente quello di cui l'utente ha bisogno: un punto unico di accesso all'informazione, dovunque questa sia localizzata. Questo può essere realizzato non solo catalogando in modo tradizionale, anche se integrato dalle altre informazioni che servono a questa particolare tipologia di documenti, ma anche, e meglio, metacatalogando. Metacatalogare significa utilizzare la metainformazione sul documento che i produttori di questo aggiungono al documento stesso. Un nuovo modo di cattura di catalogazione già fatta?

Direi proprio di sì. In questo modo i bibliotecari non devono lavorare di più per fornire questo servizio che è essenziale per gli utenti ma devono utilizzare i metadati per costruire modi di accesso privilegiati (per disciplina, per tipo di utente, per tipo di documento e quanto altro possa essere realizzato per diverse esigenze). Perché il lavoro di catalogazione dovrà cambiare diventando molto simile, addirittura difficilmente identificabile, al servizio d'informazione.

Come cambia la catalogazione

Identificare e selezionare ciò che vale è responsabilità attualmente del bibliotecario, addetto alle acquisizioni e, nel caso dei documenti elettronici, del catalogatore. Con uno scopo ben chiaro in mente e cioè essere di aiuto all'utente, che può anche essere remoto ma di cui si conoscono le esigenze. Il primo problema da affrontare è come selezionare i documenti elettronici da catalogare.

Le fonti tradizionali che vengono normalmente catalogate hanno superato una sorta ➤



di controllo di qualità, poiché sono passate attraverso vari filtri, come l'identificazione, la selezione, la raccolta, l'organizzazione, l'immagazzinamento, e l'accesso a queste fonti è ottenuto attraverso la citazione nel catalogo. Perché dovrebbe essere diverso per le fonti elettroniche? Il bibliotecario o lo specialista ha ugualmente la necessità di selezionare le risorse elettroniche e renderle accessibili attraverso il catalogo, come se facesse parte della collezione, anche se in questo caso l'accesso al documento è garantito da un programma apposito di ricerca. La selezione delle fonti elettroniche deve basarsi sugli stessi criteri di qualità del documento tradizionale. Alcuni criteri potrebbero quindi essere quello di catalogare:

- le risorse che sono importanti per la disciplina di specializzazione della biblioteca;
 - il contenuto del documento che è originale ed esaustivo su un dato argomento;
 - l'informazione che non sarebbe disponibile in altro modo;
 - i documenti che sono a testo pieno, indicizzati elettronicamente;
 - i documenti la cui versione Internet è la più nota;
 - le fonti su formati diversi (sia su carta che in rete);
 - le fonti su server che non cambiano frequentemente indirizzo o che lo comunicano quando lo cambiano;
 - le risorse che sono copiate in diversi siti. Per una risorsa importante è utile avere più di una localizzazione;
 - le risorse gratuite e non protette dal diritto d'autore oppure le risorse per cui esistano accordi con editori e fornitori di informazione;
 - le fonti di cui si conosce l'autorevolezza ed affidabilità del produttore del documento elettronico.
- Alcune scelte possibili sono:
- non catalogare i siti Web

ma solo monografie e periodici. Il problema è che spesso i documenti elettronici non hanno un titolo distinto. I periodici inoltre hanno a volte un abbonamento individuale e non istituzionale. Chi si abbona? il catalogatore o lo specialista?

— non limitare la selezione alle sole fonti elettroniche memorizzate nel server della biblioteca: a regime i soli documenti elettronici memorizzati localmente sono quelli prodotti all'interno dell'istituzione. Naturalmente non si può mai essere certi della permanenza in rete del documento. Si dovrà controllare la Url almeno tre volte, prima di eliminare il dato. L'inaccessibilità può non essere dovuta al sito ed una certa instabilità è abbastanza comune nel primo periodo della disponibilità in rete della fonte elettronica;

— le fonti elettroniche selezionate devono essere statiche nel tempo. Il problema è che spesso esistono differenti versioni dello stesso documento che non possono essere trattate come diverse edizioni dello stesso libro. Nuove versioni dello stesso documento elettronico sostituiscono di regola la precedente versione. La differenza tra le varie versioni può essere anche minima ma spesso è radicale. Le versioni possono succedersi in brevissimo spazio di tempo. Un documento elettronico è quindi incomparabilmente più dinamico di un documento a stampa. Un documento elettronico può inoltre essere rimosso senza avviso dall'accessibilità in rete o essere spostato in un altro server. Per documenti presenti su più di un server bisogna scegliere a quale fare riferimento.

Gli standard da usare non sono più gli stessi. O meglio, si è provato ad estendere il Marc, nel caso in cui i documenti elettronici in rete devono essere inseriti insieme ai documenti tradizionali nel catalogo in linea. Ma lo sforzo

che questa operazione richiede non è commisurabile con il risultato. I documenti selezionati saranno accessibili dall'opac che è il catalogo per i documenti tradizionali e esteso alle fonti elettroniche. Il problema è che non è possibile cliccare dal catalogo alla fonte su Internet.

La tendenza è che la catalogazione dei documenti elettronici sia realizzata dal fornitore dell'informazione, autore o edito-

re che esso sia. In progetti come Tulip, Ease, Decomate (Delivery of copyrighted material electronically) e Webdoc, gli editori forniscono il testo elettronico delle riviste insieme alla descrizione bibliografica, in molti casi con l'abstract. Marc è ritenuto troppo complicato ed invece vengono favoriti formati più semplici. Il Dublin Core, un formato sviluppato sotto la supervisione di Oclc e Ncsa, è il formato che molto

I formati per i metadati

Negli ultimi anni sono stati proposti molti formati per metainformazione: Capcas heads, Usmarc field 856, Dublin Metadata Core Element Set, Iafa templates, Soip, Tei Header, Urc (Uniform Resource Citation).

Usmarc field 856

Nel 1992 l'Oclc Office of Research iniziò lo studio della catalogazione delle risorse in Internet. Come conseguenza di questo progetto, il campo 856 per localizzazione ed accesso elettronico fu aggiunto a Usmarc.

Iafa templates

Questo formato è stato proposto da Internet Anonymous Ftp Archives (Iafa) che è un gruppo di lavoro dell'Internet Engineering Task Force (Ietf). Secondo questo formato, la metainformazione su Internet è presentata come un record in cui i dati hanno determinati valori e attributi. Alcuni degli elementi sono:

- Abstract e descrizione del documento
- Citazione bibliografica
- Ampiezza del file
- Lingua
- Set di caratteri

— Isbn

— Data dell'ultima revisione

Lo scopo di questo formato è l'aggiornamento di Archie, l'archivio di tutti i documenti scaricabili dalla rete Internet.

Soip (Harvest)

Il Soip (Summary Object Interchange Format) è usato dal sistema Harvest per la ricerca, recupero e l'indicizzazione dei documenti elettronici. Soip si basa in parte su Iafa.

Tei Header

Il Tei (Text encoding initiative) ha prodotto una guida per la codifica e lo scambio di testi che contiene indicazioni per la registrazione di dati catalografici usando Sgml. È usato dalla Biblioteca dell'University of Virginia. Il formato Tei ha tutti i dati che servono per una completa descrizione bibliografica insieme a tutta l'informazione documentaria non bibliografica per l'analisi del contenuto del documento, realizzata sia dal bibliotecario sia, usando i nuovi programmi di recupero dell'informazione, dalla macchina. Il formato quindi differisce dagli altri formati semplificati destinati all'uso di autori e produttori dell'infor-

probabilmente diverrà lo standard. Il Dublin Core è finalizzato proprio alla catalogazione ed indicizzazione dei documenti elettronici effettuata direttamente dai fornitori di informazione elettronica in Internet. Oclc ha sviluppato il progetto Spectrum, un programma Web adatto anche a non catalogatori per descrivere risorse elettroniche. Spectrum genera record in Usmarc, Tei (Text encoding initiative) che

è un tipo di documento in Sgml e Urc (Uniform resource citations).

La localizzazione non sarà la segnatura ma la localizzazione attraverso la quale si può accedere al documento. L'indicazione dell'Url è quella del documento e non quella del sito Web. In alternativa si può usare Urn (Uniform Resource Name) che significa nome uniforme della risorsa e che propone un identificatore stabile, unico

mazione perché è un formato complesso quanto il Marc ma differisce anche dal Marc in quanto è strettamente correlato alla presentazione del contenuto del documento (il campo Marc 55xx delle note libere non può contenere tutti i dati di Tei) ed è orientato non a descrivere una copia fisica di un documento ma a favorire il recupero analitico del suo contenuto in modo semiautomatico.

Capcas Heads (Computer aided production for current awareness services)

È il formato dei dati messo a punto da Elsevier che è la casa editrice più attiva nel campo dell'editoria elettronica già fin dal 1991. Deriva da Sgml. I dati previsti sono:

- Editore
- Periodico
- Fascicolo
- Titolo dell'articolo
- Abstract
- Parole chiave.

Urc (Uniform Resource Citation)

Lo scopo di Urc è di fare un collegamento tra la Url del documento ed il suo Urn. Ma non tutti sono d'accordo su questa funzione ed alcuni propongono di tenere autonomo Urn. In realtà la Urc assomiglia di più ad un record catalografico esteso ai dati della localizzazione in

rete. I dati previsti sono:

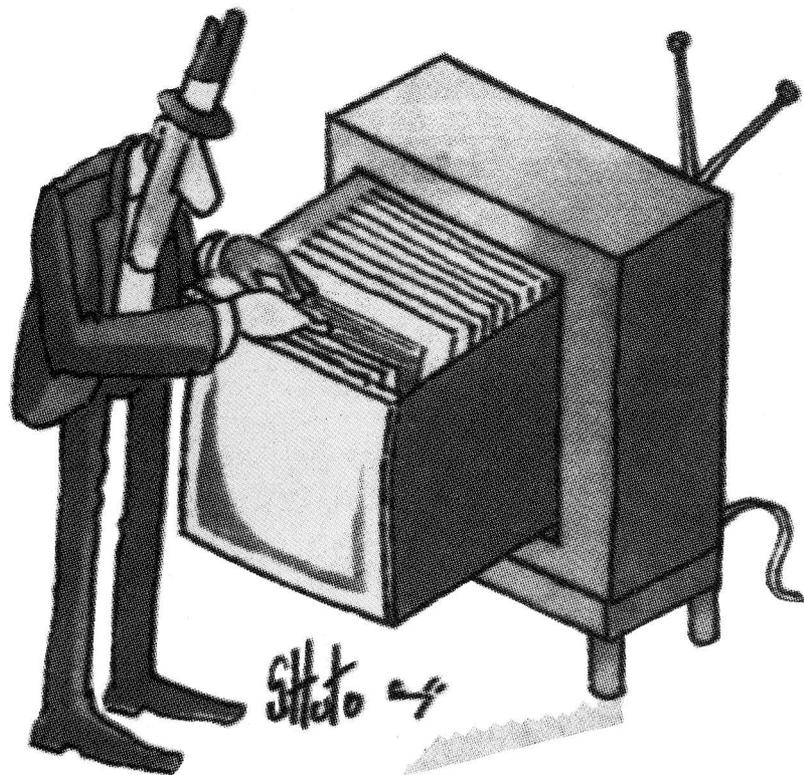
- Urn
- Autore
- Titolo
- Soggetto
- Abstract
- Url
- Firma elettronica
- Recensione
- Accesso
- Versione

La firma elettronica serve per l'autenticazione dell'autore ed è a salvaguardia della responsabilità intellettuale dell'autore della versione del documento da citare.

Urc è uno degli Uri (Uniform Resource Identifiers), identificatori univoci di risorse elettroniche. Gli Uri sono stringhe, dall'aspetto generalmente criptico per il non esperto di Internet, che hanno l'ambizione di codificare in modo semplice ed univoco una risorsa in rete. Urn è un altro identificatore, che potremmo dire simile ad Issn ed Isbn, sviluppato per la nozione di oggetto che in rete viene attribuito al documento elettronico.

Il terzo tipo di Uri è Url, la localizzazione del sito di Internet dove si trova fisicamente il documento ed è attualmente il più usato degli Uri.

Gli Uri sono usati dai motori di ricerca di Internet per l'indicizzazione automatica delle risorse in rete.



ed indipendente dal sito del documento.

La descrizione degli articoli va tenuta separata dal catalogo, in banche dati specializzate. L'utente che cerca un articolo, va direttamente ad interrogare tale archivio; solo dopo questa prima ricerca vorrà sapere se può trovare la rivista in biblioteca interrogando il catalogo. Una delle condizioni per scegliere un abbonamento elettronico, è quella che il fornitore informi con la posta elettronica che il nuovo fascicolo è disponibile. Se non c'è questo messaggio non c'è modo di sapere che il nuovo fascicolo è consultabile. Per il resto i periodici elettronici si trattano come quelli cartacei.

Alcuni progetti di metacatalogazione

Ci sono alcune interessanti iniziative bibliotecarie per catalogare le risorse di rete. Ecco di seguito una descrizione breve

dei progetti più interessanti:

Oclc Internet resources cataloging experiment è un gruppo di lavoro composto da Oclc e dalla Library of Congress. Ha studiato i problemi della catalogazione delle risorse elettroniche per i bibliotecari e prodotto un'estensione del Marc per consentire la localizzazione e l'accesso a questi documenti. il progetto *Catriona* ha realizzato diverse attività:

- catalogazione di risorse elettroniche possedute localmente e di risorse Internet rilevanti per gli interessi dell'utenza ed estensione dell'opac per ospitare i nuovi record bibliografici;
- aggiornamento del formato (simile a Marc) per aggiungere Url e Urn;
- adattamento di client standard per l'opac per consentire le funzionalità del Web di localizzazione e recupero dei documenti elettronici e funzionalità di ricerca del tipo Wais;
- costruzione di una banca dati cooperativa per la i- ➤

Dublin metadata core element set

Il Dublin Metadata Core Element Set è un formato realizzato in conseguenza di un seminario dell'Oclc e del Ncsa National Center for Supercomputing applications nel 1995 a Dublino in Ohio, da cui il formato ha derivato il suo nome. Uno degli obiettivi del seminario era quello di raggiungere il consenso su un insieme essenziale (core set of metadata elements, da cui la seconda parte del nome "Dublin metadata core element set") di elementi di metainformazione. La funzione più importante di questo insieme essenziale di dati è di facilitare il recupero di risorse in Internet effettuato dai motori di ricerca come Lycos o Webcrawler. Gli autori o i produttori di informazione, che non sono catalogatori o indicizzatori, dovrebbero essere capaci di descrivere la pubblicazione da loro prodotta usando il formato Dublin Core senza nessun addestramento.

L'insieme è formato da 13 elementi:

Soggetto	L'argomento del documento
Titolo	Il nome identificativo del documento
Autore	Il responsabile del documento
Editore	L'editore o l'agenzia responsabile della disponibilità del documento
Altri	Le persone, come ad esempio curatori e trascrittori, che hanno portato contributi intellettuali significativi all'opera
Data	Data di pubblicazione
Tipo documento	Il genere del documento come poesia, novella, ecc.
Forma	Il formato del documento come Pdf o Ms-Windows
Identificatore	Stringa o numero usato per l'identificazione univoca del documento
Relazione	Collegamento ad altri documenti
Fonte	Documento sia a stampa che elettronico da cui il documento può essere derivato
Lingua	Lingua del testo
Copertura	Copertura geografica e temporale del documento.

identificazione e la catalogazione di fonti elettroniche; — controllo ed accesso differenziato all'opac per diverse categorie di utenti.

Il progetto *Eldoradoc* tratta i documenti elettronici come quelli tradizionali a stampa. Questi sono catalogati nell'opac ed accessibili attraverso il catalogo centrale di Pica che è su

Internet. Gli utenti possono spedire i documenti che li interessano con la posta elettronica al loro indirizzo o stamparli. I problemi che il progetto ha analizzato e risolto sono:

— come organizzare il pagamento della fornitura elettronica; — come ovviare alla carenza di stampanti Postscript;

— come potenziare la capacità della rete o la memoria del Pc per chi lavora da casa;

— come realizzare la conversione di documenti che non sono in Pdf (Portable Document Format) o Postscript;

— come selezionare dei documenti liberi e non protetti dal diritto d'autore.

Il progetto cooperativo di *Bubl* tenta una classificazione e descrizione selettiva delle fonti elettroniche. Un gruppo di volontari con interessi specifici identifica e descrive le risorse rilevanti per quell'area classificate con Udc.

Il progetto *Infoservices* in Olanda, portato avanti dalla Biblioteca nazionale e da Surfnet, cominciò a catalogare secondo la tradizione archivi in un server in rete ma successivamente rinunciò all'impresa ritenuta troppo gravosa e chiese agli stessi fornitori di informazione di fornire la descrizione dei loro documenti, integrando queste descrizioni nel catalogo centrale Pica. Il progetto ha predisposto un formato semplificato ed è selettivo sulle fonti che vengono raccolte nel server.

Il progetto *Edinburgh Engineering Virtual Library* (Eevl) è specializzato nel catalogare tutte le fonti di interesse per l'ingegneria sia possedute nella biblioteca che esterne ma consultabili ugualmente con il browser standard di Web.

Per concludere, vorrei citare un'opinione su cui riflettere. Dempsey² afferma che le biblioteche, in particolare in Europa, hanno investito molti sforzi nella creazione di milioni di record bibliografici e dei cataloghi unici, ma l'insieme delle fonti disponibili è frammentato e diviso in molti canali. Quanti cataloghi dovrebbe interrogare un utente per sapere dove trovare un documento che gli interessa? Come ottenere ciò che serve con un'unica ricerca? Servono "metadati" o metacataloghi, in realtà liste di

reperimento, più che cataloghi, con la possibilità ulteriore di condurre automaticamente l'utente a visualizzare nel suo pc ciò che cerca.

Mi chiedo allora se la teoria del catalogo è ancora valida o occorra adattarla alle nuove esigenze. Oppure se quel che serve è un metacatalogo, che è un'altra cosa dal catalogo e assomiglia di più ad una banca dati con funzionalità particolari. In questo caso, lasciando invariata la struttura del catalogo tradizionale, questo potrebbe essere affiancato da metadati per collegare insieme fonti interne ed esterne. Ma per ora si va tutti per tentativi.



Note:

¹ Un esempio di un opac che integra il catalogo in linea con le fonti di rete è l'opac dell'Università di Bologna che usa un client Web standard per la ricerca sia delle fonti interne all'Università che delle fonti esterne. I cataloghi basati su Z39.50 che sono accessibili con il client standard di Web sono ricercabili alla Url: <http://lib.ncsu.edu/staff/morgan/alcuin/wwwed-catalogs.html>

Altri collegamenti a cataloghi e banche dati basati su Z39.50 sono accessibili alla Url: <http://vinca.cnidr.org/reference/reference.html>

² L. DEMPSEY, *Network resource discovery: a European library perspective in Smith N. ed. Libraries, networks and Europe: a European networking study*, London, British Library, 1994