

Raccomandazioni per miglioramenti urgenti dell'OPAC

Il ruolo delle agenzie bibliografiche nazionali

John D. Byrum, Jr.

*Library of Congress
Washington, D.C.
jbyr@loc.gov*

Vorrei iniziare facendo un'osservazione preliminare, e cioè che negli ultimi dieci anni gli utenti dell'informazione hanno incrementato considerevolmente le aspettative relative alla soddisfazione delle loro esperienze di ricerche informative come conseguenza della ricerca e dell'uso delle risorse web. Gli utenti delle bibliografie nazionali e degli OPAC di biblioteca oggi hanno aspettative di accesso all'informazione che vanno ben oltre ciò che può essere offerto da approcci di tipo tradizionale alla descrizione bibliografica. Farò un'analisi di questi sviluppi da due prospettive:

- 1) l'esigenza per i database bibliografici di garantire l'accesso ad un contenuto molto più ampio di quanto offerto tradizionalmente;
- 2) la necessità di avere una nuova generazione di OPAC che offra una funzionalità notevolmente potenziata, basata in prevalenza su caratteristiche standard dei motori di ricerca del web e delle librerie online.

Considerata la loro importanza, le agenzie bibliografiche nazionali (NBA) rivestono un ruolo privile-

giato nel processo di raggiungimento di questi obiettivi, sia grazie al potenziamento di una serie di servizi che esse stesse possono promuovere sia in base alla loro capacità di influenzare sviluppatori e venditori di sistemi integrati di biblioteca e strumenti simili.

Nel secolo scorso le NBA e altre agenzie di catalogazione hanno cercato di ottenere il controllo bibliografico e inventariale delle loro risorse tramite la produzione di record bibliografici relativi ad esse. L'avvento delle ISBD che identificano gli elementi e il loro ordine per costituire la descrizione, e delle standardizzazioni dei formati quali UNIMARC e MARC 21 ha promosso un aumento considerevole dello scambio di record. Inoltre, l'avvento di FRBR e l'elaborazione attualmente in corso di un Codice di catalogazione internazionale (ICC) costituiscono due ulteriori iniziative dell'IFLA finalizzate al raggiungimento dell'obiettivo di un Controllo bibliografico universale. Per quanto di grande valore e importanza, questi come altri successi internazionali puntano generalmente alla creazione e allo scambio delle descrizioni bibliografiche di tipo tradizionale, e mirano prevalentemente alle caratteristiche di base delle risorse di biblioteca – in altre parole, le descrizioni si limitano essenzialmente agli elementi dei dati obbligatori stabiliti dalle ISBD.¹ Gli utenti di oggi sono sempre più



John D. Byrum, Jr.

condizionati dalle loro esperienze nel web ad avere accesso all'informazione in modo rapido e facile e ad apprezzarne i risultati, cosa che li aiuta a stabilire l'utilità dell'informazione ricevuta. Nello scenario della bibliografia nazionale e del catalogo di biblioteca, queste aspettative sottolineano l'esigenza di arricchimenti bibliografici che di solito non vengono forniti. Eppure, grazie all'automazione e ad altri sviluppi ad essa collegati, una certa gamma di informazioni potrebbe diventare in modo facile e poco dispendioso una caratteristica costante dei nostri database bibliografici. Mi spiego meglio portando ad esempio l'esperienza della Library of Congress (LC), dove abbiamo implementato con successo

Per gentile concessione dell'autore, pubblichiamo in queste pagine la traduzione dell'importante intervento tenuto da John D. Byrum, Jr. alla 71. IFLA Conference, che si è svolta a Oslo il 14-18 agosto 2005.
Traduzione di Grazia Di Bartolomeo

una serie di progetti con l'obiettivo di collegare ricercatori, cataloghi e risorse web, ampliare il contenuto del record catalografico in sé (ad esempio, includendo i TOC-Tables of Content e le recensioni) e collegare le descrizioni bibliografiche alle relative risorse elettroniche (ad esempio, l'accesso via web alle pubblicazioni individuali all'interno delle collezioni). Nel 1992 il responsabile della catalogazione della Library of Congress ha dato vita al Bibliographic Enrichment Advisory Team (BEAT), composto interamente da unità di personale proveniente a titolo volontario da una serie di uffici con lo scopo di fare ricerca e avviare iniziative finalizzate all'arricchimento dei record bibliografici. Il gruppo è stato incoraggiato a intraprendere un tipo di lavoro "fuori dagli schemi" ed è stato esentato dalle limitazioni imposte dalle politiche e dalle pratiche pertinenti alle attività di catalogazione di tipo tradizionale. Il BEAT ha avuto l'incarico di approntare e usare metodi automatizzati per raggiungere i propri obiettivi e di avere un impatto minimo sulle attività dello staff responsabile della catalogazione tradizionale.²

Uno dei primi sforzi del BEAT si è concentrato sull'arricchimento dei record bibliografici, includendo informazioni relative ai TOC, perché da ricerche effettuate si è arrivati alla conclusione che tali informazioni migliorano notevolmente la funzione di recupero del database e sotto molti aspetti l'esperienza dell'utente con il catalogo. Ad esempio, alcuni studi effettuati negli anni Ottanta e Novanta hanno portato alla conclusione che i TOC aiutano gli utenti a determinare meglio l'importanza di alcuni titoli ai fini delle proprie esigenze informative, che le parole chiave dei TOC migliorano notevolmente l'efficacia della ricerca e che i TOC offrono il mezzo per superare le

ristrettezze tradizionali della ricerca per soggetto. In effetti, uno studio durato undici anni ha stabilito che "la ricerca per soggetto [è] stata rimpiazzata dalla ricerca per parole chiave", un punto decisamente a favore dei TOC, in quanto essi sono interamente costituiti da parole chiave.³

Poiché la biblioteca non poteva permettersi che lo staff inserisse manualmente i dati TOC, la sfida portata avanti dal BEAT è stata quella di trovare una soluzione informatica per ottenere l'obiettivo dell'arricchimento. In effetti, il BEAT ha fornito una serie di strumenti e di programmi. Il primo è stato la creazione di un'applicazione chiamata "Text Capture and Electronic Conversion" (TCEC) per la catalogazione delle pubblicazioni che la LC riceve in formato elettronico dagli editori che partecipano al programma Cataloging-in-Publication (CIP). Usando il software TCEC, il catalogatore mette in evidenza i TOC a partire dal manoscritto elettronico. Il programma quindi elabora i dati TOC e aggiunge il risultato nel campo 505 del record catalografico (che nel MARC21 è destinato all'informazione dell'indice). L'elaborazione effettuata da TCEC comprende la formattazione dell'informazione relativa ai contenuti sulla base delle specifiche delle AACR2 per la registrazione dei dati TOC – ad esempio, cancellare i termini e la numerazione di capitolo, sezione, parte; eliminare la paginazione e aggiungere la punteggiatura delle ISBD. TCEC converte anche tutte le parole in minuscolo, tranne la prima parola del titolo di ogni capitolo, così il catalogatore non deve fare altro che mettere in evidenza i nomi propri che devono essere scritti in maiuscolo. Meglio ancora, il trasferimento dell'informazione dal manoscritto al record si verifica nel giro di pochi secondi e i dati vengono registrati in

modo preciso così come compaiono sul manoscritto elettronico. Ad oggi sono state create un totale di oltre 40.000 note di TOC usando TCEC.⁴ Il numero è destinato a crescere, in quanto stanno aumentando le schede CIP in formato elettronico presentate dagli editori. Solo quest'anno sono stati inviati per la catalogazione qualcosa come 35.000 manoscritti elettronici; il 35-40 per cento delle relative descrizioni bibliografiche è stato ritenuto adeguato per l'inclusione dei dati TOC.

Per questo progetto e per altre iniziative relative ai TOC sono stati adottati tre criteri principali per la selezione delle pubblicazioni ai fini dell'elaborazione:

- 1) i titoli dei capitoli devono contribuire a migliorare la ricerca per parola chiave tramite linguaggio naturale;
- 2) i titoli dei capitoli devono garantire una maggiore comprensione dei contenuti dell'articolo rispetto a quanto trasmesso dal titolo e dall'indicazione dell'area di responsabilità.
- 3) i dati TOC non devono richiedere un gravoso lavoro manuale di editing per essere disponibili per l'elaborazione elettronica.

Fortunatamente quasi tutti sono in grado di applicare questi criteri nel momento in cui vengono prese delle decisioni rapide. Altri tre progetti sui TOC si basano ampiamente su applicazioni software. Il primo è il progetto Digital Tables of Contents (D-TOC), avviato alla fine degli anni Novanta. Questa iniziativa fa affidamento sul personale, prevalentemente tecnici all'inizio della carriera e studenti borsisti di scuole superiori, che passano allo scanner i TOC di pubblicazioni scelte secondo criteri predefiniti, come quelli citati prima. A quel punto entra in gioco un'applicazione software: i TOC passati allo scanner vengono convertiti in testo usando l'OCR, poi codificati

in html e trasferiti su uno dei server della LC. Durante questa conversione vengono modificati automaticamente anche i record catalografici originari in formato MARC al fine di includere i collegamenti con i dati TOC, rendendo in questo modo reciproco il collegamento tra le due fonti d'informazione. Per migliorare la possibilità di recupero dell'informazione, le parole del record bibliografico appartenenti ai campi del titolo e dell'indicazione di responsabilità vengono presentate all'inizio della visualizzazione dei TOC. Inoltre, il tag dei metadati della parola chiave nel file TOC html contiene le parole dei termini di accesso per soggetto al record bibliografico. Il TOC presente nel web offre un collegamento diretto al record originario del catalogo, e a sua volta il record del catalogo fornisce un collegamento reciproco dal campo 856 al record TOC nel web. Inoltre, poiché Yahoo!, Google e altri motori di ricerca indicizzano questi record D-TOC, gli utenti di Internet non solo vengono a conoscenza dei record html, ma hanno anche la possibilità di accedere all'OPAC della LC tramite i collegamenti presenti in questi record. Un fatto interessante è che dal 6 giugno 2005 una ricerca fatta con Yahoo! sulla frase "contenuti per il controllo della Library of Congress" produce un risultato di 262.000 record, tutti collegati con i D-TOC.

Fino ad oggi, come risultato di questo progetto, sono stati inseriti nel web circa 30.000 TOC, con una percentuale attualmente di circa 300 la settimana. Di recente il progetto D-TOC è stato ampliato per poter includere alcune collezioni retrospettive al fine di dare una copertura delle pubblicazioni in altre lingue, a partire dal tedesco. La produzione dei D-TOC dovrebbe presto aumentare ulteriormente, perché la biblioteca ha di-



Paul Klee, *Via principale e vie traverse*, 1929

versi uffici dislocati oltreoceano (Il Cairo, Nuova Delhi, Islamabad, Rio, Nairobi, Giacarta) dove si sta lavorando per coinvolgere queste sedi nella creazione di file TOC per poi trasmetterli alla LC.

Il più produttivo e al tempo stesso meno costoso tra i progetti TOC del BEAT è l'applicazione ONIX-TOC avviata nel 2000. Questa iniziativa consiste nell'estrazione dei dati TOC dai file ONIX forniti dagli editori. Il progetto si basa su un programma Visual Basic che passa allo scanner una serie di file ONIX per creare degli indici digitali che

vengono poi messi a disposizione nel web. Come per il progetto D-TOC, ciascuno dei record ONIX-TOC offre all'utente la possibilità di visionare il record bibliografico nel catalogo online della LC per avere ulteriori informazioni. Allo stesso modo, i record bibliografici dell'OPAC della LC vengono arricchiti automaticamente con collegamenti nel campo 856 dei file ONIX-TOC. Alcuni record ONIX-TOC includono anche le immagini delle copertine. Attualmente, sono stati creati oltre 50.000 di questi record, che sommati ai D-TOC a lu-

glio 2005 davano un totale di oltre sei milioni di titoli.

Il limite più significativo dei progetti D-TOC e ONIX-TOC è la loro scarsa influenza sulla capacità di recupero dei record bibliografici nell'OPAC, poiché i record prendono la forma di dati disponibili sul web. All'inizio di quest'anno, però, si è registrato un notevole passo avanti grazie allo sviluppo di una procedura informatica che permette di analizzare ed estrarre automaticamente i dati necessari per la struttura del testo TOC dai record TOC digitali già esistenti nel web. Questa informazione viene poi modificata e resa conforme al formato richiesto dal sistema integrato della biblioteca, consentendo così di aggiungere il TOC al record catalografico. Poiché questi dati TOC non sono pienamente conformi alle specifiche delle AACR2, essi vengono presentati con la annotazione: "Nota di contenuto generata automaticamente". Questa tecnica verrà adottata per il futuro, inoltre lo staff del BEAT sta provvedendo a elaborare i file dei 30.000 record D-TOC creati in passato. Si verificherà anche se lo stesso tipo di approccio può essere applicato agli ONIX-TOC, che nel frattempo sono accessibili solo tramite i motori di ricerca del web e tramite collegamenti nel record bibliografico nell'OPAC.

Sono sceso nel dettaglio per quanto riguarda i progetti di arricchimento dei contenuti bibliografici al fine di mettere in evidenza il ruolo che l'automazione gioca nell'arricchire il contenuto delle descrizioni bibliografiche. Nessuna di queste iniziative ha un costo in termini di tempo impiegato dal personale, una volta che i programmi vengono scritti e messi a punto. In termini, invece, di spese effettive di operazione, le cifre sono sorprendentemente basse. Ad esempio, il costo medio globale per la creazione di un record ONIX-TOC si aggi-

ra intorno a 0,10 centesimi di dollaro o anche meno ciascuno.

I TOC sono solo una delle attività di arricchimento attualmente in corso alla LC. Ne citerò qualche altra senza scendere troppo nel dettaglio. In primo luogo c'è una serie di progetti basati su dati ONIX disponibili gratuitamente, oltre a quelli utilizzati per gli indici. Una delle applicazioni ONIX che la LC ha avviato offre agli utenti del web e dell'OPAC la possibilità d'accesso alle sintesi fatte dagli editori per dare un'indicazione del contenuto dei libri da loro pubblicati. Lo staff BEAT fa una valutazione iniziale di queste sintesi prima di decidere di elaborarle. Per esempio, vengono respinte le descrizioni di un particolare editore che sembrano eccessivamente promozionali. Invece, ogniqualvolta una fonte di tali sintesi viene accettata, il record arricchito con i collegamenti ad esse porta una nota che dice: "L'informazione proviene da dati elettronici forniti dall'editore", così gli utenti vengono informati del fatto che la fonte delle sintesi non è la LC. Ad oggi, più di 100.000 record sono stati aggiornati per consentire l'accesso a queste descrizioni.

In un altro progetto ONIX i file di alcuni estratti campione forniti dall'editore (di solito uno o più capitoli, oppure la prefazione) vengono resi disponibili sul web con l'aggiunta dei collegamenti ai relativi record bibliografici. Qualche capitolo campione è disponibile in html, mentre altri sono in formato .pdf, il che, nei casi in cui ci sono illustrazioni, rende possibile agli utenti vederne gli ingrandimenti. Finora sono stati elaborati oltre 17.000 estratti campione.

Infine, a partire dallo scorso anno, le informazioni biografiche fornite dagli editori relative ai collaboratori delle pubblicazioni selezionate sono state rese disponibili in rete e collegate con i record bibliografici.

Ad oggi sono stati elaborati oltre 12.600 di questi arricchimenti.

In un altro settore dell'attività di arricchimento bibliografico, il gruppo BEAT ha dato vita a una serie di accordi che hanno portato alla possibilità di accedere alle note scientifiche e alle recensioni per alcune particolari acquisizioni. In alcuni casi, come nei progetti TOC, l'accesso è all'informazione che viene catturata e resa disponibile nel web, con i collegamenti aggiunti al relativo record bibliografico. Il criterio principale per poter decidere quale tecnica utilizzare è la lunghezza del testo. Nel caso di recensioni più brevi, il programma automaticamente estrae l'informazione e la inserisce nei corrispondenti record bibliografici. Al contrario, quando le recensioni sono lunghe e includono le note e le bibliografie, vengono rese disponibili in rete creando i collegamenti ad esse nelle relative descrizioni bibliografiche. In ogni caso, la biblioteca richiede il permesso di poter mantenere una copia delle recensioni utilizzate in uno dei nostri server, al fine di assicurarne la longevità. Le fonti del progetto relativo alle note e alle recensioni devono rispettare dei criteri particolari. Per poter essere selezionate, le recensioni devono essere: disponibili gratuitamente, costanti nella pubblicazione, di carattere scientifico, con una responsabilità chiara, e in file leggibili dal programma. Poiché ci sono svariate agenzie che pubblicano recensioni a pagamento, questo progetto del BEAT ha avuto modo di crescere meno rapidamente rispetto ad altri progetti di arricchimento. Comunque, finora sono stati selezionati sette collaboratori, tra cui:

- 1) "fonti di riferimento eccellenti", dalla Reference and User Services Association dell'American Library Association;
- 2) siti selezionati dal Best Free Reference Web Sites Committee della

Machine-Assisted Reference Section di questa associazione;

3) *Handbook of Latin American studies*, pubblicato dalla Library of Congress;

4) "H-Net reviews in the Humanities and Social Sciences", rivista di carattere scientifico disponibile online;

5) *Library resources and technical services*, pubblicato dall'Association for Library Collections and Technical Services;

6) "Education Review", dell'Arizona State University.

7) "Bryn Mawr Classical Reviews". Altri due progetti del BEAT che si basano anch'essi prevalentemente su applicazioni software dimostrano come la bibliografia nazionale e altre agenzie possono arricchire l'accesso collegando le descrizioni bibliografiche direttamente alle risorse elettroniche full-text. Il primo è "Web access to publications in series", che ha come obiettivo quello di aumentare l'utilità della letteratura grigia di un certo valore nei settori della finanza e dell'economia. A maggio del 2005 questo progetto aveva prodotto i collegamenti per l'accesso a diverse migliaia di titoli individuali contenuti in circa 305 collezioni di natura prevalentemente attinente al campo della ricerca e dello sviluppo. Di queste collezioni, circa il 40 per cento in precedenza non erano rappresentate nell'OPAC della LC. Per consentire la creazione di record bibliografici individuali per le monografie individuali in queste collezioni pubblicate elettronicamente è stato creato un programma denominato Web Cataloging Assistant, che crea record MARC estraendo ed elaborando i dati direttamente dai titoli monografici all'interno della collezione selezionata per l'analisi. Usando questo strumento un tecnico consulta la pagina dell'abstract per la monografia e incolla la URL di quella pagina nel programma. Web Cataloging Assistant poi recupera la

pagina web e cattura il testo da quella pagina relativamente al titolo, agli autori, al numero di serie, alla sintesi e alle parole chiave, se disponibili. Il programma aggiunge anche altre informazioni di base in forma di note standard. Web Cataloging Assistant infine aggiunge questi record nel database della LC. Il programma informa anche lo staff con comunicazioni automatiche via posta elettronica quando del materiale nuovo viene aggiunto nei siti web che le società e simili creano per le loro pubblicazioni. Da giugno 2004 a luglio 2005 sono stati creati quasi 3.000 record usando Web Cataloging Assistant. Questa applicazione non solo ha messo la biblioteca in condizioni di fornire accesso alla letteratura grigia in tempo sempre più rapido, ma di farlo a una frazione del costo normale persino per un record a livello minimo.

L'ultimo progetto del BEAT che descriverò è "Web access to works in the public domain," il cui obiettivo è quello di consentire il recupero del full text delle versioni elettroniche delle monografie a stampa della collezione della LC. Questo progetto è il risultato di diverse collaborazioni per mezzo delle quali alcuni depositi di fiducia forniscono alla LC un elenco elettronico di URL e di numeri di controllo della Library of Congress di pubblicazioni monografiche che essi hanno digitalizzato a uso del pubblico. Un'applicazione software recepisce questa informazione nella biblioteca, identifica i record bibliografici corretti posseduti dalla Library of Congress per la versione a stampa, e aggiunge poi le URL per le versioni elettroniche. Il risultato è che all'utente viene dato immediato accesso alla versione elettronica e che quindi non ha bisogno di recuperare la versione a stampa. La biblioteca partner viene menzionata in una nota che identifica e quindi pubblicizza la

collezione in cui si trova la versione elettronica.

I progetti sull'arricchimento bibliografico che ho descritto finora dimostrano come, nell'era dell'elettronica, una biblioteca nazionale prenda in considerazione dei servizi tradizionali, e li arricchisca con nuove modalità per investire su di essi nell'era digitale. Questi progetti sono facili da realizzare e poco costosi da gestire, richiedono un modesto numero di persone dedicate, per lo più per sviluppare le applicazioni software citate. Una volta avviati, i programmi funzionano per la maggior parte dietro le quinte in modo autonomo. Molte agenzie bibliografiche nazionali hanno la competenza necessaria per rispondere a questa funzionalità, e così facendo possono aprire le porte ai loro database, ai loro OPAC e al loro posseduto. Attività come quelle descritte offrono una risposta agli utenti che sono sempre più abituati ai risultati in tempo reale delle ricerche per parola chiave e che chiedono qualcosa di più dell'informazione bibliografica di base. Fornendo questi arricchimenti le biblioteche nazionali si mettono al servizio delle accresciute esigenze di informazione del ricercatore, offrono percorsi strutturati alle loro fonti di informazione, e quindi invogliano gli utenti a seguire ad apprezzare i loro prodotti e i loro servizi bibliografici. Gli arricchimenti mettono anche in condizione gli OPAC di funzionare al meglio come meccanismi di accesso soddisfacenti.

Sulla base di questa riflessione, passiamo adesso al secondo argomento che ho proposto di trattare, e cioè l'esigenza di una nuova generazione di OPAC con una funzionalità notevolmente arricchita, che può in gran parte basarsi sulle caratteristiche standard dei motori di ricerca del web e delle librerie online. Per cominciare dall'aspetto positivo, prendiamo atto del fatto

che gli OPAC di oggi effettuano nel complesso una vasta gamma di compiti in modo abbastanza efficace. In effetti, uno dei più grandiosi progressi della tecnologia dell'informazione nel XX secolo è stata la trasformazione del catalogo cartaceo in catalogo in linea ad accesso pubblico, fulcro dei sistemi integrati di biblioteca (ILS) di oggi. Il Codice di catalogazione dell'IFLA, in via di elaborazione, identifica cinque funzioni fondamentali di questo strumento – cercare, identificare, selezionare, ottenere e navigare – e in molti conveniamo sul fatto che la generazione attuale di OPAC esegue queste funzioni in modo soddisfacente.

Sempre più spesso, comunque, in letteratura e altrove vengono espresse preoccupazioni sul fatto che i nostri cataloghi in linea siano vicini alla loro maturità, e che, in particolare, non siano adeguati per il recupero delle risorse digitali in crescente aumento. Mentre dappertutto le riviste elettroniche diventano una componente sempre più presente nelle biblioteche nazionali e di ricerca, gli OPAC recuperano facilmente l'informazione bibliografica sui titoli, ma generalmente non consentono all'utente di accedere direttamente ai singoli articoli in esse contenuti. Le NBA e le biblioteche in generale non possono permettersi di sviluppare le descrizioni bibliografiche al di sotto di questo livello. Inoltre, a causa delle ristrettezze economiche, gli OPAC forniscono al massimo l'accesso solo a una selezione relativamente piccola di risorse elettroniche accessibili da postazioni remote, compresi i database di citazioni, aggregazioni di full text, strumenti di reference online e siti web, perché queste risorse sono in genere voluminose e pesanti, di una certa complessità, tanto da sfuggire al controllo bibliografico universale e aggiornato. Sempre più di frequente, gli uten-

ti non si affidano agli OPAC per raggiungere le risorse di loro interesse.⁵ Piuttosto, la stragrande maggioranza si è prontamente rivolta a motori di ricerca veloci e spesso precisi quali Google, che si sono sviluppati in risposta alla proliferazione dei materiali in rete. In effetti, sembra indiscutibile che questa grande maggioranza sia soddisfatta di tali strumenti, anche se spesso incontra set di risultati voluminosi e poco precisi.⁶ Intorno al 2000, comunque, sono cominciati a emergere nuovi sistemi di biblioteca pensati per rispondere a esigenze informative più sottili, alcuni nostrani e altri di tipo più commerciale. Tra questi strumenti ci sono motori di ricerca federati (a volte denominati "portali"), OpenURL resolver e sistemi di gestione delle risorse elettroniche (ERMS), tutti aventi il fine di connettere l'utente con le risorse elettroniche, comprese quelle disponibili all'interno di una collezione di biblioteca ma non accessibili per mezzo dell'OPAC o quelle accessibili da remoto nel web.⁷

Tutti insieme, questi nuovi strumenti cercano di offrire una funzionalità che consenta una singola ricerca di un elevato numero di database bibliografici di alto livello e di risorse full-text, comprese informazioni approfondite nel web che conducono gli utenti al miglior materiale e forniscono un tipo di recupero esauriente per un determinato argomento o compito – con la facilità d'uso di un motore di ricerca come Google. Sebbene le versioni iniziali di questi strumenti fossero chiaramente al di sotto delle aspettative, quasi tutte le grandi biblioteche ne hanno acquistato una serie, in quanto offrono agli utenti un accesso immediato all'informazione stessa, e non semplici descrizioni bibliografiche di materiale di interesse affine. In realtà gli OpenURL resolver, le applicazioni di ricerca meta-federate

e gli ERMS si sono dimostrati così utili, e la loro richiesta così elevata, che la competizione ha stimolato i venditori a investire generosamente per poter offrire il prodotto più ricercato dopo averlo sottoposto ad ulteriori miglioramenti.

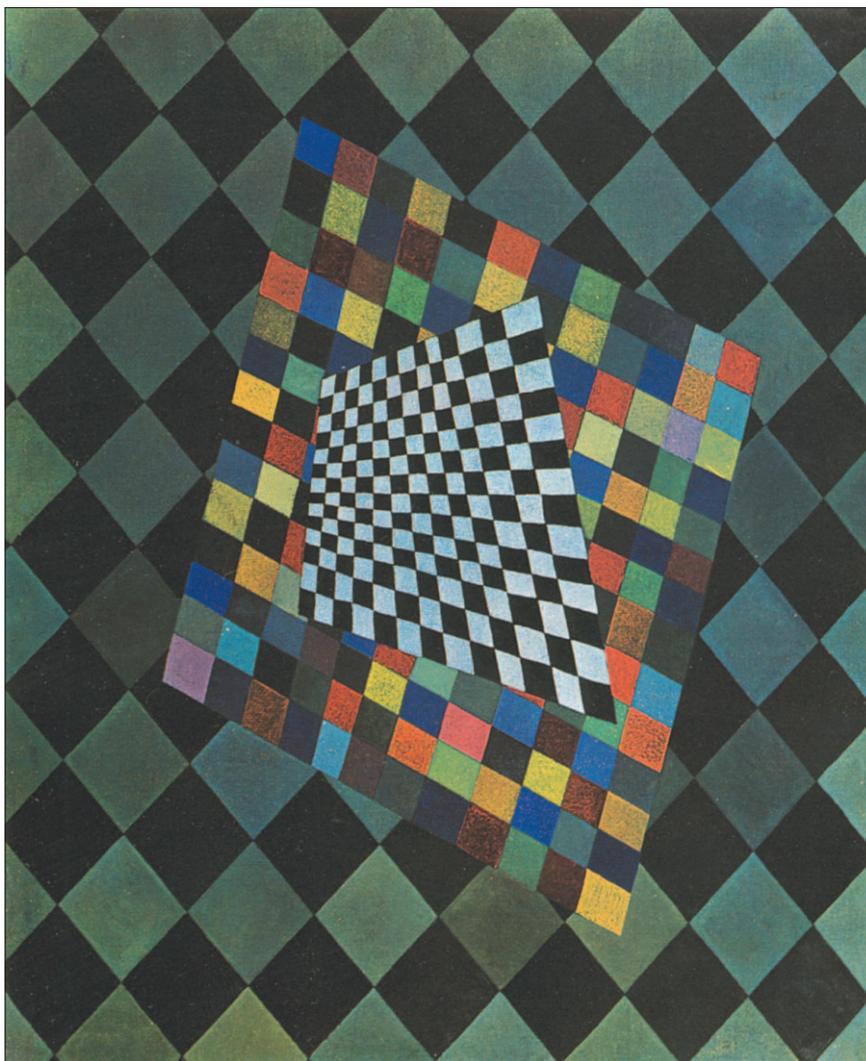
Nonostante l'avvento e l'arricchimento di questi strumenti indispensabili, per la maggior parte essi rimangono al di fuori del dominio dell'OPAC. Così, nelle sue osservazioni introduttive a *Il futuro dei sistemi integrati di biblioteca: una tavola rotonda*, Brian Kenney sostiene:

La potenzialità di migliorare l'interoperabilità spinge i bibliotecari a guardare in modo più critico agli ILS. Alcuni si domandano se il portale delle informazioni del futuro debba essere basato su un solo ILS o se invece debba essere una serie di prodotti di diversi venditori di ILS. Questo tipo di approccio all'interoperabilità nel mercato della biblioteca, un po' per volta, è stato creato dagli stessi venditori con alcuni prodotti come SFX di Ex Libris, strumento usato per il collegamento delle fonti di riferimento, e ENCompass di Endeavor, un prodotto nato per la creazione e la gestione dei contenuti digitali. I prodotti per il collegamento e la gestione del digitale hanno costituito, da soli, circa il 13 per cento del mercato degli ILS dello scorso anno.⁸

Questa osservazione rimane valida oggi come lo era nel 2003, quando fu scritta. La Library of Congress è interessata ad esplorare la possibilità di sviluppo di una strategia che miri a ottenere un migliore allineamento della funzionalità dell'OPAC con le caratteristiche di questi nuovi strumenti. Come risultato, la LC sta finanziando una ricerca che analizzi l'evoluzione della natura del catalogo in linea, compresa la possibilità di inquadrare il tutto in un contesto di integrazione con questi altri strumenti di ricerca. Questa iniziativa offrirà anche

l'opportunità di esplorare un'aspettativa di Sarah Thomas che già dal 2000 prefigurava un'eventuale "reinterpretazione" del catalogo nella direzione di "un servizio di informazione che registra in modo sistematico quelle pubblicazioni e quei documenti di interesse per una particolare comunità, a prescindere dal loro formato". La Thomas anticipò anche l'avvento a breve termine di un "ibrido, che avrebbe preso alcune caratteristiche migliori del catalogo, ma che avrebbe utilizzato un'infrastruttura tecnologica sofisticata per incrementare la resa nei confronti degli utenti dell'informazione".⁹ Il progetto di ricerca della LC, concepito in collaborazione con numerosi esperti, ha l'obiettivo di identificare cosa può essere fatto ora per orientarsi verso tale obiettivo; le raccomandazioni che ne scaturiranno, e che saranno ampiamente condivise, indicheranno fasi successive di natura pratica.

Oltre alla necessità di integrazione della funzionalità, la nuova generazione di OPAC dovrà necessariamente offrire molte altre peculiarità, se i servizi delle NBA e i cataloghi vorranno competere con le migliori applicazioni di oggi per la ricerca delle risorse, rappresentate dai motori di ricerca commerciali e dalle librerie online. Nel loro studio su *The impact of Web search engines on subject searching in OPAC*,¹⁰ Yu e Young hanno riscontrato che "l'uso dei motori di ricerca del web è sufficientemente presente ovunque, tanto da avere un impatto sulle aspettative degli utenti di un OPAC basato sul web". La loro ricerca li ha portati a sostenere la necessità di una serie di caratteristiche per le interfacce della prossima generazione di OPAC basati sul web, in modo da produrre dei sistemi per gli utenti "che verosimilmente applicano al catalogo di biblioteca un modello mentale tipico dei motori di ricerca del web". Le funzionalità



Vasilij Kandinskij, *Quadrato*, 1927

specifiche raccomandate da Yu e Young sono:

– *Sequenza dei menu*. Alcuni studi hanno dimostrato che gli utenti scelgono le opzioni di ricerca posizionate sull'estremità più alta della lista, quindi bisognerebbe dare la giusta considerazione a quegli approcci posti in alto. Thorne e Whitlach sostengono che i menu dovrebbero incoraggiare fortemente i neofiti a iniziare con ricerche per parola chiave. Un'analisi delle connessioni delle transazioni effettuata presso la California State University di Los Angeles ha messo in evidenza il fatto che il successo nella ricerca migliora quando l'interfaccia è stata ridisegnata per

offrire la seguente sequenza: parola/frase chiave, titolo, autore, soggetto (specificando che è richiesta la Library of Congress Subject Heading) e indice di collocazione.

– *Caratteristiche di browsing*. Hancock-Beaulieu ha riscontrato che il 30-45 per cento delle ricerche online viene seguito da un'attività di browsing fra gli scaffali della biblioteca. Secondo Yu e Young "una opzione di browsing può aiutare gli utenti a reperire in modo efficace i documenti desiderati raggruppando documenti affini per argomento sulla base dei termini di un thesaurus".

– *Visualizzazione dei risultati*. L'impatto della ricerca nel web sugli utenti ha indotto questi ad

aspettarsi rappresentazioni grafiche, come le icone, per la visualizzazione dell'informazione. Le visualizzazioni dell'OPAC dovrebbero essere riprogettate per incorporare la grafica che rende più facile la lettura dello schermo e la comprensione delle definizioni.

– *Ordini di pertinenza.* Gli utenti dei motori di ricerca del web dipendono da questi per conferire un ordine di priorità ai set di risultati ottenuti. Raramente viene visualizzata più di qualche pagina di riferimenti. Yu e Young suggeriscono l'uso di algoritmi di ordini di pertinenza che includano data di pubblicazione, intestazioni per soggetto e termini affini, aggiungendo che “la popolarità fra gli utenti e la frequenza dei termini negli indici” potrebbero essere altri fattori impiegati.

– *Fornitura di suggerimenti utili.* Gli OPAC dovrebbero poter offrire aiuto senza che gli utenti lo richiedano, mettendo a disposizione strumenti e concetti che diano la possibilità di attuare strategie di ricerca migliori. In aggiunta, i sistemi dovrebbero generare al volo suggerimenti sulla ricerca e, in caso di risultato zero, suggerire parole alternative.

Yu e Young sostengono che la nuova generazione di OPAC debba avere incorporate anche altre caratteristiche offerte di frequente dai motori di ricerca e dalle librerie online, in quanto conferiscono un valore aggiunto all'esperienza di chi cerca l'informazione. Sempre più spesso la letteratura scientifica registra articoli di diversi altri autori che concordano con queste raccomandazioni.

– *La capacità di correzione dello spelling come la funzione di Google “forse cercavi”.* Yu e Young suggeriscono che errori di spelling nelle transazioni effettuate al catalogo potrebbero essere usate come base per consentire agli OPAC di offrire questo servizio.

– *Tenere traccia della popolarità fra gli utenti, come in Amazon.com.* Agli utenti del catalogo verrebbe offerta la possibilità di esprimere i loro commenti sulle scelte, che funzioni come una sorta di “sistema di raccomandazione”; i loro commenti, insieme alle correlazioni relative alla circolazione, potrebbero funzionare come base per avvisare gli utenti sulla disponibilità di pubblicazioni affini.

– *Ricerca tramite linguaggio naturale come nella funzione “singola casella di ricerca” di Google.* Gli utenti del catalogo dovrebbero avere la possibilità di accedere ai titoli tramite le parole iniziali degli articoli e i nomi degli autori posti in qualsiasi ordine, proprio come fanno quando usano i motori di ricerca.

Di recente la Library of Congress ha incaricato Marcia Bates di effettuare uno studio avente l'obiettivo di esplorare le modalità di arricchimento dei record di metadati fornendo ulteriori meccanismi di accesso quali quelli per soggetto e altri (ad esempio, i thesauri di consultazione immediata per l'utente) e maggiore sottigliezza di accesso e visualizzazione. La relazione finale di Bates propone una serie di raccomandazioni per incrementare l'accesso dell'utente all'informazione proveniente dal catalogo di biblioteca e dal portale.¹¹ Volendo semplificare i contenuti della sua ricerca, Bates sostiene il principio che riconoscere è di gran lunga più affidabile che richiamare, in quanto è una caratteristica propria della ricerca e che un thesaurus di consultazione immediata metterebbe gli utenti in condizioni di usare termini diversi quando ci si appresta a fare la ricerca di un determinato argomento e di ottenere i risultati desiderati. La Bates ha quindi sviluppato un approccio concettuale per offrire agli utenti un vocabolario con struttura a grappolo in risposta alla parola chiave o alle parole digitate nel si-

stema. L'approccio prevede la creazione di “un sistema di thesaurus per l'utente” che sia progettato come possibilità di consultazione immediata rispetto ai database bibliografici esistenti. La sua proposta dettagliata è sicuramente di grande interesse per i venditori che intendono migliorare la prossima generazione di OPAC. Nel frattempo, la stessa LC sta valutando se è in grado di implementare alcuni aspetti dello studio della Bates per i propri scopi.

Ci conforta il fatto che ci siano già in corso iniziative che confermano la necessità di migliorare l'esperienza dell'utente con l'OPAC. Negli Stati Uniti, per esempio, i maggiori esponenti delle agenzie che offrono servizi bibliografici stanno mettendo a punto una serie di iniziative per accrescere il valore dei loro cataloghi online. Impressionante per gli sforzi che sta compiendo per rafforzare il catalogo è il progetto “RedLightGreen” di RLG, sostenuto dalla Mellon Foundation, che ha come obiettivo quello “di offrire un'informazione di biblioteca ricca e affidabile, unica nello scenario del web, e di trasmettere tale informazione nelle modalità che soddisfano le aspettative degli utenti accorti del web”. RedLightGreen è nato per andare incontro alle esigenze degli studenti universitari di primo livello. Secondo quanto affermato da Richard Parker della University di Warwick (UK): “[Con RedLightGreen] è facile avere dei risultati, è facile rifinire la ricerca, e puoi fare molto con quanto hai ottenuto”. Seguono una serie di esempi che lui fa:

– è chiaro come procedere alla ricerca per parole/frasi chiave;
– l'utente può rifinire i risultati usando le intestazioni per soggetto della LCSH;
– i risultati vengono ordinati secondo due criteri: importanza in base al/ai termine/i di ricerca; e

numero di biblioteche con i posseduti; le risorse in cima alla lista sono possedute da più biblioteche rispetto a quelle più in basso nella lista stessa. Quindi, più in alto è collocata nell'elenco, maggiore è la credibilità accademica di una risorsa (e più probabile è che sia posseduta da una certa biblioteca);

- tutte le edizioni di un testo vengono collocate in un unico risultato, così è facile distinguere un testo letterario popolare da opere secondarie su di esso;

- gli utenti possono verificare se la loro biblioteca ha la particolare risorsa che essi stanno cercando;

- i risultati vengono messi in collegamento con altre risorse online, come articoli e recensioni, e con Amazon.com o altre librerie. Infine,
- gli utenti di RedLightGreen possono produrre bibliografie online, formattandole in uno dei quattro stili più popolari, possono inoltre scaricarle e inviarle per posta elettronica.¹²

È significativo il fatto che RedLightGreen implementi il quadro di riferimento dell'IFLA dei Functional Requirements of Bibliographic Records (FRBR), in quanto esso aggrega opere per le quali c'è un vasto numero di edizioni in un set gestibile di risultati che risponde ai termini di ricerca dell'utente. Informazioni relative agli sforzi in corso per accrescere la funzionalità di RedLightGreen, e commenti sulla sua validità, sono disponibili sul sito Web di RLG.¹³

In modo analogo, OCLC sta concentrando i suoi sforzi per migliorare l'obiettivo e l'utilità del suo catalogo online collettivo tramite una serie di iniziative come ad esempio l'Open WorldCat per "integrare i record di biblioteca con siti di ricerca Internet popolari e verificare l'efficacia del web nel guidare gli utenti verso i materiali posseduti dalle biblioteche". FictionFinder, un'altra delle iniziative di OCLC, è un sistema prototipo

sviluppato per un totale di oltre due milioni e mezzo di record bibliografici di narrativa, raggruppati al livello dell'opera in modo tale da mostrare informazioni relative a sintesi, genere, ambientazione e soggetto. FictionFinder funziona anche come prototipo per un'implementazione più ampia dei concetti di FRBR sui quali si basa. Tramite un progetto denominato da OCLC "Curiouser", l'applicazione verrà estesa ai record di Open WorldCat che Yahoo! sta raccogliendo. Inoltre, OCLC ha programmato per la fine di quest'anno di mostrare in FirstSearch i record raggruppati secondo i principi di FRBR.

OCLC ha annunciato che dall'inizio di quest'estate il sito web di Open WorldCat offre informazioni sugli strumenti che sono di aiuto per una ricerca più appropriata di materiali di biblioteca direttamente dal proprio pc, tra cui:

- 1) Yahoo! Toolbar, un'edizione speciale di Yahoo! Toolbar che dà possibilità di accesso immediato ai record Open WorldCat tramite Yahoo! Search;

- 2) Google Toolbar con la funzione Autolink, che dà anch'essa accesso immediato ai record Open WorldCat;
- 3) le estensioni di ricerca Firefox, dalle quali è possibile digitare i termini di ricerca e selezionare il motore di ricerca da interrogare.¹⁴

Ci sono progetti in corso per la creazione di siti di tipo Wiki in WorldCat. Secondo quanto afferma uno degli sviluppatori:

L'idea è quella di avere un sito Wiki che integri WorldCat. La gente potrebbe contribuire con delle recensioni, decorazioni, commenti ecc. e collegarli ai record bibliografici... Speriamo che il sistema sia sufficientemente flessibile in modo da mettere le persone in condizioni di fare delle (buone) cose che noi non ci aspettiamo. Ci piacerebbe che Wiki fosse disponibile ovunque siano i record WorldCat.¹⁵

Questi esempi dovrebbero essere sufficienti per incoraggiare le NBA ad avviare dei processi di miglioramento che rendano la navigazione nei loro database online più confacente agli utenti, le cui esperienze di recupero dell'informazione sono condizionate dalle esperienze nel web. In alcuni casi le NBA, come RLG e OCLC, troveranno conveniente investire al loro interno per fornire quell'incremento di funzionalità raccomandato. In altri casi, le NBA potrebbero esercitare la loro influenza, individualmente oppure in uno sforzo condiviso, per richiedere una migliore funzionalità ai venditori che commercializzano e curano la manutenzione degli OPAC da loro usati. Come affermato senza mezzi termini da Roy Tennant in un recente articolo comparso su "Library Journal":

Dobbiamo concentrare maggiormente le nostre energie su quei cambiamenti sistematici importanti piuttosto che su quelli estetici. Se un sistema di ricerca è più difficile e meno efficace di Amazon.com (e quale non lo è?), allora dobbiamo lavorarci su. Finiamola di richiedere sottigliezze ai venditori. Dopo tutto, certo che possiamo mettere il rossetto a un maiale, ma resta pur sempre un maiale.¹⁶

Le nostre bibliografie nazionali e i nostri cataloghi sono il risultato di secoli di sforzi intellettuali e di risorse investite. Considerati gli enormi investimenti portati avanti per creare e mantenere queste interfacce nelle loro collezioni, ovunque le biblioteche dovrebbero cercare l'opportunità di fornire quelle funzionalità necessarie per invogliare gli utenti a continuare ad usare gli OPAC come sistemi di accesso soddisfacenti. Esercitando la loro autorevolezza, le NBA rivestono un ruolo privilegiato per adottare modalità nuove e migliorate con l'obiettivo di investire sui

servizi tradizionali nell'era digitale tramite attività di arricchimento delle risorse bibliografiche, e sostenendo la necessità di una urgente revisione dell'OPAC per meglio servire le esigenze degli utenti del XXI secolo. È giunto il tempo per un'azione immediata, affinché i nostri utenti non abbandonino non solo le risorse che abbiamo raccolto per loro, ma anche gli strumenti che le rendono disponibili.

Note

¹ Ulteriori informazioni relative alle attività dell'IFLA sono disponibili nella Cataloguing Section della home page all'indirizzo: <<http://www.ifla.org/VII/s13/index.htm>> (ultima consultazione: giugno 2005).

² Ulteriori informazioni relative al Bibliographic Enrichment Advisory Team e a tutti i progetti descritti di seguito sono disponibili nella home page del BEAT all'indirizzo: <<http://www.loc.gov/catdir/beat/>> (ultima consultazione: giugno 2005).

³ Alcuni studi e relazioni che convalidano l'importanza dei TOC sono: EVAN PAPPAS – ANN HERENDEEN, *Enhancing bibliographic records with table of contents derived from OCR technologies at the American Museum of Natural History Library*, "Cataloging and Classification Quarterly", 23 (2000), 4, p. 65-67; R. CONRAD WINKLE, *An analysis of table of contents in recent english-language books*, "Library Resources and Technical Services", 43 (1998), 1, p. 14; RUTH C. MORRIS, *Online tables of contents for books: effect on usage*, "Bulletin of the Medical Library Association", 89 (2001), 1, p. 29. Si rimanda anche alla home page di RichCat, all'indirizzo: <<http://www.loc.gov/standards/catenrich/>> (ultima consultazione: giugno 2005).

⁴ In un secondo progetto denominato E-CIP TOC, facendo uso di programmi elaborati da tre membri del BEAT sono stati creati dei TOC basati sul web per virtualmente il 100% dei record E-CIP che contengono dati TOC. Questi dati sono generati dal programma e per ogni voce viene fatto

un collegamento diretto nei TOC verso e da il record originario che si trova nel catalogo della LC. I programmi (modificati di recente) gestiscono ora la maggior parte dei segni diacritici, e arricchiscono anche la visualizzazione dei TOC sul web con le intestazioni per soggetto della LC attribuite dallo staff della catalogazione. Da maggio 2005 sono stati aggiunti sul server del web circa 54.000 record Electronic CIP (E-CIP) TOC.

⁵ MARSHALL BREEDING, *The many facets of managing electronic resources*, "Computers in libraries", 24 (2004), 1; <<http://www.infotoday.com/cilmag/jan04/brbreeding.shtml>>.

⁶ JOEL ACHENBACH, *Search for tomorrow: we wanted answers, and Google really clicked, what's next?*, "Washington Post", 15.02.2004, D7.

⁷ Una panoramica esauriente di questi strumenti è disponibile presso il sito del Portals Applications Issues Group della Library of Congress all'indirizzo: <<http://www.loc.gov/catdir/lcpaig/paig.html>> (ultima consultazione: giugno 2005).

⁸ BRIAN KENNY, *The future of integrated library systems: an IJ round table*, "Library Journal", 15.06.2003, 37.

⁹ SARAH E. THOMAS, *The catalog as portal to the Internet*, in *Proceedings of the Bicentennial Conference on bibliographic control for the new millennium: confronting the challenges of networked resources on the web*, Washington, D.C., Cataloging Distribu-

tion Services, 2001, 35; <http://lcweb.loc.gov/catdir/bibcontrol/thomas_paper.html> (ultima consultazione: giugno 2005).

¹⁰ HOLLY YU – MARGO YOUNG, *The impact of Web search engines on subject searching in OPAC*, "Information Technology and Libraries", (Dec. 2004), p. 168-180.

¹¹ MARCIA BATES, *Task force recommendation 2.3 Research and Design Review... final report (Version 3), June 1, 2003*, disponibile all'indirizzo: <<http://www.loc.gov/catdir/bibcontrol/2.3BatesReport6-03.doc.pdf>> (2005).

¹² RICHARD PARKER, *Promoting Red-LightGreen at the University of Warwick*, "RLG Focus", (Dec. 2004), 71; <http://www.rlg.org/en/page.php?Page_ID=20480#article0> (ultima consultazione: giugno 2005).

¹³ <http://www.rlg.org/en/page.php?Page_ID=13061> e <http://www.rlg.org/en/page.php?Page_ID=20500> (ultima consultazione: giugno 2005).

¹⁴ OCLC abstracts, 8:23 (June 6, 2005), <<http://www5.oclc.org/downloads/design/abstracts/06062005/index.htm>> (ultima consultazione: giugno 2005).

¹⁵ THOM HICKEY, *Library metadata techniques and trends*, blog inviato nel maggio 2005, <http://outgoing.typepad.com/outgoing/2005/05/worldcat_wiki.html> (ultima consultazione: giugno 2005).

¹⁶ ROY TENNANT, *Lipstick on a pig*, "Library Journal", 15.04.2005, p. 34.

Abstract

Today's information seekers have been conditioned by Web search engines to expect immediate gratification as the result of user-friendly Web experiences. In contrast, it is increasingly apparent that traditional library OPACs do not provide the same ease of use or access to information. National Bibliographic Agencies (NBAs) and libraries everywhere need to respond to this discrepancy by initiating measures to enrich their databases and bibliographic products with much more information than is currently captured in records for resources. At the same time, NBAs must address the need for a new generation of OPACs that offers significantly enhanced functionality, much of which can be based on standard features of Web search engines and online bookstores. In view of alternatives available to information seekers, these needs require immediate attention if NBAs and libraries are to retain the support of satisfied users into the 21st century. This paper offers specific recommendations to assist them in identifying and implementing appropriate responses.