

# Di che parla questo catalogo?

*Un'indagine sugli accessi semantici  
negli OPAC italiani*

Claudio Gnoli

*Dipartimento di matematica  
Università degli studi di Pavia  
gnoli@aib.it*

Riccardo Ridi

*Dipartimento di studi storici  
Università "Ca' Foscari" di Venezia  
ridi@aib.it*

Giulia Visintin

*Biblioteca consulente  
visintin@aib.it*

Organizzare i documenti in base al loro contenuto sembra essere un'operazione piuttosto di moda grazie a Internet: l'enorme massa di informazioni in forma digitale che progressivamente divengono disponibili produce una situazione caotica, e per gestirla con maggiore efficienza numerosi esperti di informatica invocano l'uso di metadati relativi ai contenuti informativi. Come è noto lo stesso creatore del World Wide Web, Tim Berners-Lee, ha auspicato una sua evoluzione in direzione di un "Web semantico", nel quale gli elementi di contenuto siano adeguatamente marcati e collegabili con informazioni affini attraverso schemi standard.<sup>1</sup>

Da un punto di vista teorico, queste idee non si differenziano molto dalle tradizionali tecniche di catalogazione semantica sviluppate da decenni in biblioteconomia e documentazione; purtroppo però l'insufficiente integrazione tra le comunità di bibliotecari e informatici e l'uso di terminologie completamente diverse producono un alto tasso di "reinvenzioni dell'acqua calda", con grande spreco di risorse e di occasioni applicative. Questa situazione suggerisce di fare un passo indietro, guardando ai tradizionali cataloghi delle biblioteche come al luogo primario in cui ci si può aspettare di vedere applicato il grande patrimonio di tecniche d'indicizzazione consegnatoci dalla ricerca passata.

Si è pensato perciò di andare a verificare se, nell'incombente era del Web semantico, siano già semantici almeno i nostri OPAC, consapevoli che anche in questo caso l'esistenza di tecniche avanzate non garantisce di per sé una loro applicazione realmente efficace. Quali possibilità di ricerca semantica (per soggetti, per termini di tesauri, per classi) offrono i cataloghi in rete italiani ai loro utenti? Avremmo bisogno di buoni esempi di OPAC semantici anche per mostrare ai colleghi informatici il funzionamento e l'utilità delle tecniche d'indicizzazione che abbiamo sviluppato con il lavoro di oltre un secolo... Già diverse voci tuttavia si sono levate a lamentare la scarsa applicazione ai cataloghi automatizzati dei potenti strumenti a disposizione dei bibliotecari.

Dopo aver per qualche tempo coltivato l'idea di un repertorio di OPAC che rendessero disponibili una ricerca classificata, si è poi optato per un'indagine più sistematica, grazie anche alla disponibilità a collaborare da parte di un gruppo di studenti dell'Università Ca' Foscari di Venezia. Si è inoltre cercata la collaborazione di tutti i soggetti interessati a lavorare in questa direzione, in modo particolare del Gruppo di ricerca sull'indicizzazione per soggetto (GRIS) dell'AIB, autorevole voce che da tempo sostiene l'uso di criteri coerenti e standardizzati nell'indicizzazione

semantica. Ai coordinatori del progetto si è quindi affiancato un gruppo aperto di rilevatori volontari e di collaboratori, mantenuti in contatto fra di loro attraverso la posta elettronica e il web.<sup>2</sup>

Il progetto si è articolato in diverse fasi. Un necessario lavoro preliminare è stata la messa a punto di una lista di controllo, da applicare poi ai cataloghi considerati. La scelta delle voci e la loro formulazione richiedeva particolare attenzione, in quanto queste avrebbero poi costituito un riferimento stabile e possibilmente non più modificato. Mentre alcune voci sono apparse immediatamente scontate (la presenza delle principali funzioni di ricerca per soggetti e per classi e la loro presentazione nei risultati), su altre ci si è dovuti soffermare in dettaglio: ad esempio la definizione esatta della ricerca in due fasi per termine e stringa, indicata dai membri del GRIS come una funzionalità cruciale, e la definizione delle voci riguardanti i sistemi di indicizzazione usati nel catalogo (quali soggetti, quali tesauri, quali schemi di classificazione e in quali edizioni); per queste ultime la pratica ha imposto che il rilevatore indicasse il sistema usato solo nei casi (purtroppo minoritari) in cui esso fosse dichiarato esplicitamente nel catalogo, rinunciando invece a qualsiasi tentativo di inferirlo dall'aspetto delle schede, operazione che avrebbe comportato rischiose ambiguità.

### Nota sul metodo

Gli OPAC italiani censiti dal repertorio AIB-CILEA al momento dell'indagine erano oltre 600: perciò, per considerarne un campione significativo, occorre avvalersi di un buon numero di rilevatori.

Il nucleo principale di essi è stato costituito dai 13 studenti del corso di Documentazione elettronica tenuto da Riccardo Ridi dell'Università Ca' Foscari di Venezia, corso che nell'anno accademico 2003-2004 verteva sulla interrogazione degli OPAC: Samantha Biolcati, Federica Busetto, Serena Cairo, Laura Carnelos, Antonella Daulerio, Fiamma De Salvo, Claudia Favaron, Silvia Girardello, Chiara Guidolin, Laura Paccagnella, Alessia Savio, Francesca Tedeschi, Diego Vedoato. A loro si sono aggiunte, in qualità di collaboratrici volontarie, altre quattro bibliotecarie, tre delle quali provenienti dal master in Scienze dell'informazione di Padova Ricerche: Lorena Zuccolo, Emanuela Casson e Beatrice Motta con la collaborazione di Vera Dean.

*Il criterio di scelta degli OPAC* – È stato individuato nel livello di copertura geografica riportato dal repertorio: sono stati considerati gli OPAC a copertura provinciale (ossia comprendenti biblioteche situate in più di un comune), regionale (biblioteche di diverse province) e nazionale (biblioteche di diverse regioni), tralasciando così quelli relativi a biblioteche di uno stesso comune o a una sola biblioteca, indipendentemente dalla lo-

ro importanza. 19 OPAC comunali del Veneto e del Friuli-Venezia Giulia (area di provenienza dei rilevatori) sono comunque stati inclusi, anche per sostituire OPAC a più ampia copertura ai quali il collegamento in rete è risultato ripetutamente impossibile nel periodo dei rilevamenti.

Nell'ottobre 2003 a Venezia si è tenuto un seminario, nel quale Claudio Gnoli e Riccardo Ridi hanno istruito i rilevatori in merito al significato della ricerca semantica, sulle modalità operative e sui criteri del rilevamento. Ogni rilevatore ha preso in carico una decina di OPAC, senza particolari criteri di suddivisione del campione. Naturalmente l'impiego di una squadra di rilevatori richiede che si presti la massima attenzione all'omogeneità dei criteri di valutazione e all'interpretazione dell'esatto significato delle voci della lista e dei codici con i quali vengono registrati. A questo proposito, oltre agli accordi presi di persona durante l'incontro e le successive lezioni del corso, è stata regolarmente utilizzata una lista di discussione, tramite la quale gli studenti erano in costante contatto con i coordinatori del progetto, per chiarire qualsiasi dubbio o ambiguità. L'utilizzo di questo mezzo di comunicazione, utile per rafforzare il senso di appartenenza a una comunità di lavoro telematica, ha costituito una delle parti più peculiari e a nostro avviso interessanti dell'esperienza di ricerca. Gli studenti, dopo qualche in-

certezza e timore iniziali, hanno dimostrato di impiegare tale strumento con una certa naturalezza, anche per accordarsi fra di loro su qualche occasione di incontro diretto per confrontare il lavoro. Si è così potuto effettivamente realizzare quel connubio fra didattica e ricerca che in teoria dovrebbe essere connaturato all'insegnamento universitario, specialmente nelle lauree specialistiche, ma che troppo spesso resta solo teorico. I dati sono stati registrati da ciascun rilevatore su un foglio di calcolo appositamente predisposto con una suddivisione standard delle colonne, e quindi consegnati ai coordinatori che li hanno ricomposti in un file unico. È stato poi svolto un ulteriore attento lavoro di omogeneizzazione dei dati, in particolare per alcune voci risultate ostiche ed estranee all'esperienza degli studenti, quale quella relativa alla fantomatica ricerca per termine e stringa. Il foglio dei dati risultanti è stato pubblicato sul sito del progetto e rimane disponibile per chiunque fosse interessato, o che volesse informarsi su aspetti particolari, eventualmente anche contattando i coordinatori. Un'analisi statistica descrittiva ha quindi permesso di sintetizzare i risultati per ciascuna voce in forma di tabelle, anch'esse pubblicate sul sito. Il presente articolo, insieme ad appositi incontri pubblici, costituisce la presentazione dei risultati e delle principali considerazioni che da essi emergono.

Parallelamente, è stata avviata una raccolta di riferimenti bibliografici sull'argomento, in particolare degli ultimi due decenni, utilizzando come base di partenza BIB per i contributi in italiano e LISA per quelli in altre lingue. Sono stati individuati numerosi articoli che hanno trattato il problema della ricerca per soggetto negli OPAC, in alcuni casi anche con analoghe indagini e con esperimenti sul campo con gli u-

tenti, prevalentemente in area anglosassone. La bibliografia, pubblicata sul già citato sito del progetto, rimane aperta a successivi incrementi, in modo da offrire uno strumento aggiornato.

Nell'autunno 2003 sono stati svolti i rilevamenti principali, applicando la lista di controllo a un ampio campione di OPAC. In seguito i risultati sono stati analizzati e riassunti in forma statistica: tutti i dati

sia analitici che raggruppati sono pubblicamente disponibili sul sito del progetto. Nel seguito ci concentriamo sulla descrizione dei risultati più significativi.

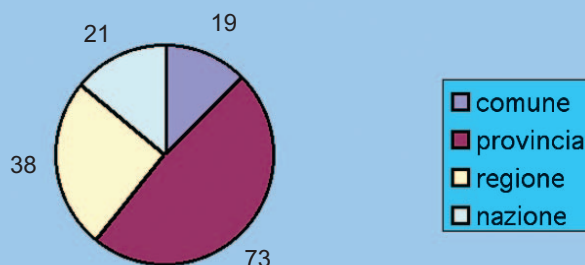
#### **Informazioni sugli accessi semantici**

Questo aspetto dei cataloghi è risultato uno dei più carenti, no-

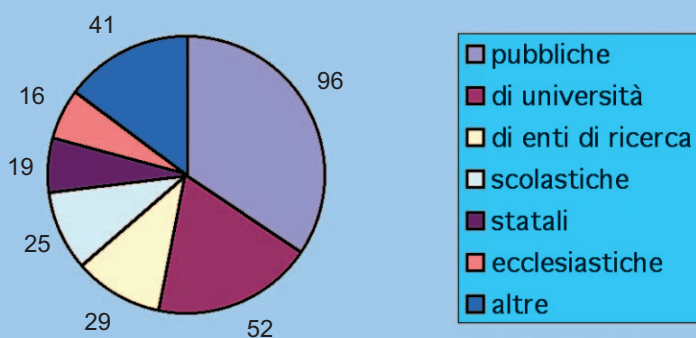
nostante richiede per essere curato pochi sforzi e nessun requisito tecnico. Ben pochi OPAC infatti si premurano di informare gli utenti sull'uso di sistemi di organizzazione semantica e sulla loro identità, limitandosi eventualmente a presentare campi di ricerca per soggetti o classi, come se l'opzione proposta fosse l'unica possibile.

I cataloghi esaminati sono stati 152 (per la scelta del campione, vedi box nella pagina a fianco): ben 136 di essi non indicano quale parte del loro patrimonio sia indicizzata tramite soggetti o per classi, e nessuno specifica se per caso faccia uso di tesauri. Una maschera di ricerca che era accompagnata da buone indicazioni in questo senso, quella dell'Università di Modena e Reggio Emilia, è stata successivamente sostituita da una in cui le indicazioni non ci sono più. Poco migliore la situazione per quanto riguarda l'identità dei sistemi usati: in circa un terzo dei cataloghi è possibile capire di quali sistemi si tratta (perlopiù il *Soggettario per i cataloghi delle biblioteche italiane* e la CDD, ma anche altri; solo l'OPAC denominato *Bibliografia internazionale leonardiana* fa esplicito riferimento all'uso delle norme GRIS); in molti casi si tratta di cataloghi che utilizzano il software Sebina, in cui un menù a tendina permette di scorrere le liste di voci appartenenti a uno specifico sistema quando ne siano usati più d'uno: questo è però possibile solo a partire dalla prima pagina di risposte, mentre all'inizio di ciascuna ricerca l'utente è ancora virtualmente ignaro di tale possibilità. Perlopiù i sistemi usati sono nominati nelle istruzioni sul significato dei campi o dei codici per la ricerca libera, e non indicati come informazione di per sé. Interessante il fatto che, tra le liste che fanno riferimento a sistemi speciali, compaiano quattro volte un non meglio specificato

### Area di copertura dei cataloghi



### Cataloghi suddivisi per tipo di biblioteca



*Soggettario immagini* e tre volte *Iconclass*.<sup>3</sup>

Solo in 5 cataloghi è possibile sapere quale edizione della CDD sia impiegata, solo in 3 se le collocazioni siano formate con numeri di classe (sì in 2 casi su 3; l'ICTP di Trieste indica esplicitamente che è usata la CDU), e nessun catalogo si preoccupa di indicare in quale lingua siano espressi gli equivalenti verbali.

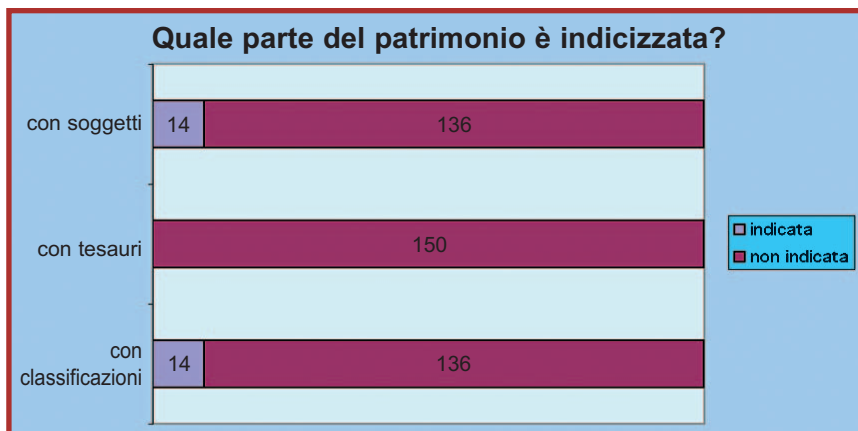
Non molti i cataloghi che offrono tesauri (7 casi) o schemi di classificazione (36 casi) a supporto della ricerca semantica; la maggioranza di questi utilizza Sebina, messo in commercio con pagine originali di spiegazioni piuttosto ricche sulle ricerche per soggetti e per classi, con particolare riferimento alla CDD; di solito questo testo è lasciato inalterato nei diversi cataloghi, senza adattarlo alla situazione specifica delle biblioteche comprese.

### Accesso per soggetti

L'accesso per voci di soggetto è piuttosto diffuso: esso è presente nel 75% delle maschere di ricerca per liste e nell'85% delle maschere di ricerca per campi elementare. Il fatto che in nessun caso sia presente nella ricerca per campi avanzata *ma non* in quella elementare suggerisce che esso sia sempre considerato uno dei canali di ricerca principali. In qualche caso i termini cercati appartengono a un vero e proprio tesaurus; ad esempio la Rete di donne in Toscana offre una ricerca per "parole chiave". In altri cataloghi, tuttavia, quest'ultima espressione è usata ambiguamente per significare invece parole presenti in qualsiasi campo, mescolando al significato documentario di *termine* quello, diverso, informatico; questa ambiguità si riscontra peraltro anche all'estero, ad

esempio viene lamentata da Will.<sup>4</sup> Generalmente, inserendo delle parole nel campo *soggetto* si ottengono in risposta direttamente le schede che le contengono in una qualunque posizione di quel campo. Il GRIS auspica invece<sup>5</sup> una ricerca per termine e stringa, che in un primo passaggio offra una lista dei termini controllati disponibili (ciascun termine potendo essere costituito da una sola o da più parole). Scelto un termine, si rende visibile la lista delle stringhe di soggetto che lo contengono: stringhe a loro volta collegate alle rispettive schede bibliografiche. Una tale ricerca offre all'utente maggior controllo e consapevolezza della forma dei termini d'indicizzazione. La lettura delle stringhe di soggetto garantisce inoltre un'informazione sul contesto nel quale il concetto rappresentato dal singolo termine è trattato in ciascun documento.

Purtroppo nessuno dei cataloghi esaminati è risultato disporre di questa funzione, e solo 5 – che utilizzano il software Zetesis e quello messo a punto presso l'Università di Firenze – ne offrono una molto simile, imperniata però su qualsiasi parola (facilmente estraibile per via automatica), o addirittura su qualsiasi frammento di parola, invece che su termini controllati. In vari OPAC Sebina, UOL e Unibiblio si possono cercare i singoli descrittori che compongono i soggetti, ma poi da questi si arriva direttamente alle schede senza passare per le stringhe, forse nell'intento di evitare troppi passaggi agli utenti. In 17 casi è possibile limitare la ricerca per soggetti a un solo vocabolario controllato (ad esempio, "Soggettario immagini" e non "BNCF"); si tratta perlopiù di OPAC Sebina, nei quali però la scelta appare solo dopo una prima ricerca. Offrono invece la scelta fin dalla prima maschera il Sistema bibliotecario ticinese e il Politecnico di Torino, quest'ultimo celan-



dola però sotto la criptica alternativa "soggetti" e "soggetti formali".

### Accesso per classe

La ricerca per classi è disponibile un po' meno spesso di quella per soggetti: nel 54% delle maschere per liste e nel 51% di quelle per campi elementari, mentre in 20 casi su 35 è presente nella ricerca per campi avanzata e non in quella elementare. Curioso il caso dell'ICTP, in cui la ricerca per classe non porta a una lista di documenti, bensì a una mappa della biblioteca in cui viene messo in evidenza lo scaffale in cui si trovano i volumi della classe cercata.

Spesso, anche se le maschere non prevedono un esplicito accesso per classe, è possibile eseguire tale ricerca inserendo i codici di classe in un campo per "ricerca libera" – funzione chiaramente utile soprattutto ai bibliotecari già esperti del sistema: ciò avviene ad esempio nella maggior parte degli OPAC Easyweb. È evidente quindi che i dati di classificazione, anche quando sono presenti, non sono considerati utili come canale di ricerca, venendo piuttosto visualizzati in fase di risposta come una sorta di informazione descrittiva. Ciò è confermato dal fatto che solo il 20% dei cataloghi permette di cercare per classi attraverso gli equivalenti verbali, che sono il punto di par-

tenza più naturale per un utente, mentre molto più numerosi sono quelli che richiedono di conoscere già il numero di classe di cui si ha bisogno. La situazione può spiegarsi con il notevole impegno necessario a mantenere aggiornate negli OPAC le tabelle degli equivalenti verbali, in particolare dei numeri costruiti in fase di indicizzazione e non presenti nelle tavole a stampa,<sup>6</sup> oltre al fatto che la sola parte liberamente riproducibile della CDD sono i simboli numerici.<sup>7</sup> Tuttavia, dal punto di vista degli OPAC, è una grave mancanza che priva la classificazione di molte delle sue potenziali applicazioni.

Una funzione particolarmente appropriata ad essere applicata ai numeri decimali di classe è quella di troncamento a destra del simbolo. Infatti in 30 casi (Sebina, Playlib, Bookmarkweb, alcuni Aleph), oltre alle schede con il numero di classe cercato, vengono mostrate automaticamente anche quelle con numeri di classe che iniziano per le cifre cercate, mentre in altri 41 è possibile attivare a richiesta questo tipo di risposta. La possibilità di effettuare il troncamento a sinistra o all'interno è invece rara (15 casi) e mai evidenziata, anche se sarebbe utile, ad esempio per le suddivisioni tratte dalle tavole ausiliarie della CDD come -03 (dizionari), -05 (periodici), -45511 (Firenze). Quando siano usati più sistemi di classificazione, solo in un terzo dei



casi è possibile limitare la ricerca a uno solo di essi; particolarmente chiare in questo senso le maschere di Politecnico di Torino e Regione Lombardia.

### Visualizzazione in fase di ricerca per liste

La possibilità di scorrere le liste di singole parole contenute nel campo *soggetto* è offerta da 41 cataloghi, compresi i 5 che offrono la ricerca per parola e stringa di cui si è detto sopra. 112 visualizzano liste di stringhe di soggetto che cominciano con le parole cercate, ma solo 39 offrono le stringhe che le contengono anche in altre posizioni, funzione che invece è molto importante per un recupero completo e non casuale;<sup>8</sup> negli OPAC Easyweb si può ottenere solo in seconda battuta, attivando la voce non molto evidente “espandi la ricerca per termini della lista”.

Per quanto riguarda la ricerca per classi, gli equivalenti verbali sono elencati raramente (16%), solo in 4 casi è disponibile una lista delle singole parole che li compongono, e solo in 5 casi sono elencati gli equivalenti verbali che contengono la parola cercata in qualsiasi posizione. Si distinguono in queste funzioni i due OPAC che utilizzano il software sviluppato dall'Università di Firenze, particolarmente chiari ed efficaci; anche gli OPAC Bookmarkweb mostrano le parole presenti in qualsiasi posizione. Molto più frequenti le liste di numeri di classe: 35% senza equivalenti verbali e 17% con equivalenti verbali, sempre o almeno in parte. L'eventuale presenza di rinvii è stata rilevata scorrendo un campione di 50 voci consecutive. I rinvii dovrebbero essere una componente assolutamente normale dei cataloghi, ma purtroppo nella maggioranza degli OPAC sono risultati assenti o presenti solo sporadica-

mente: 6 cataloghi offrono rinvii da parola a stringa di soggetto che la contiene, 11 offrono relazioni tesauroali da termine a termine, 4 da stringa a stringa – nessun catalogo offre rinvii da classe a classe. I pochi casi positivi offrono però esempi di come l'integrazione del tesoro nel catalogo possa produrre realizzazioni efficaci: il Polo SBN di Romagna, per esempio, visualizza chiaramente i collegamenti dei termini cercati con *termine associato* e *termine più ampio*, e gli OPAC Aleph di Università di Genova ed ENEA visualizzano in una finestra a parte (forse con qualche problema di accessibilità) una scheda con termine preferito, numero identificativo e termini non preferiti: ad esempio, cercando *case* si ha la gradita sorpresa di scoprire che si tratta anche di un acronimo per il termine preferito *ingegneria del software assistita dall'elaboratore*. Sorprendente l'OPAC Playlib del Sistema bibliotecario della Brianza comasca: la ricerca col termine *case* produce automaticamente e tacitamente anche le voci che iniziano coi termini *abitazioni* e *architettura*...

### Visualizzazione nelle schede

La forma sintetica delle schede ottenute in risposta a una ricerca comprende quasi sempre solo i dati bibliografici descrittivi: quelli semantici sono riservati eventualmente alle schede dettagliate. Anche in queste ultime, la presenza degli equivalenti verbali delle classi è purtroppo molto meno frequente (28%) di quella dei numeri di classe (73%), dimostrando anche qui una scarsa attenzione alla funzione della classificazione per gli utenti. Talvolta gli equivalenti verbali sono visualizzati solo per una parte delle classi, ad esempio nel catalogo della Provincia di Brescia solo per le classi dello schema di classificazione locale.

In genere, da ciascuna voce di soggetto, termine di tesoro o numero di classe presente nel tracciato di una scheda è attivabile un rinvio che permette di cercare altri documenti indicizzati allo stesso modo. Gli OPAC Aleph delle Università di Genova e di Siena si distinguono per offrire anche in questa fase l'alternativa tra una ricerca per campi (che altrimenti è la norma) o per liste. Talvolta invece la possibilità di questo tipo di rinvio non è sfruttata: l'informazione è solo visualizzata, senza essere utilizzabile come punto di partenza per ulteriori esplorazioni. In particolare, per i numeri di classe si sono registrati ben 25 casi di questo tipo, e addirittura su 43 cataloghi che visualizzano gli equivalenti verbali, solo quello del Sistema bibliotecario del Vimeratese permette di usarli per recuperare gli altri documenti classificati allo stesso modo.

### Indice di semanticità

Le voci della lista di controllo sono state impostate e codificate in modo tale che un maggiore numero di risposte positive corrisponda a una maggiore offerta di accessi semantici da parte dell'OPAC. Così facendo è stato possibile anche definire un indice di semanticità del catalogo, variabile potenzialmente tra 0 e 100, cosicché l'OPAC semantico ideale dovrebbe totalizzare 100 punti. Si tratta chiaramente di un valore approssimativo, tuttavia va precisato che la formula che calcola l'indice include una pesatura dell'importanza relativa attribuita alle diverse funzioni di ricerca e di visualizzazione, di modo che il valore complessivo rappresenti in maniera bilanciata le prestazioni dell'OPAC nei diversi settori.<sup>9</sup>

Gli indici ottenuti variano tra 54,7 e 0,0, con una media di 26,5, mostrando che l'indicizzazione semantica è generalmente lontana dal-

Tab. 1 - I primi ventun cataloghi con miglior indice di semanticità

	Indice	OPAC	Software	Copertura*	Biblioteche**
1	54,7	SBN. Polo di Romagna	Sebina	R	E N P S
2	51,8	SBN. Polo bolognese (compresa università)	UniFI	R	N P U Z
3		Università di Padova (SBN. Polo universitario veneto)	UniFI	R	E N P R U
4	46,4	SBN. Poli Sapienza + Lincei + capitolino + giuridico	Sebina	R	R U Z
5	45,3	Provincia di Pesaro e Urbino + SBN. Polo PU	Sebina	P	P S R U Z
6		SBN. Polo bolognese (compresa università). 2	Sebina	R	N P U Z
7	44,4	Provincia di Piacenza	Sebina	P	P S Z
8		Provincia di Ancona (SBN. Polo di Jesi)	Sebina	P	P
9	44,2	SBN. Polo Università di Macerata	Sebina	R	N P U
10	43,9	SBN. Polo del Friuli Venezia Giulia	Sebina	R	E N P R U
11	43,6	Provincia di Mantova	Sebina	P	P
12	42,5	SB parmense (compresa Università di Parma)	Sebina	P	E N P R U
13	42,2	Provincia di Brescia	Sebina	P	P Z
14		Provincia di Caltanissetta	Sebina	P	P
15	41,7	Soprintendenza ai beni culturali ed ambientali di Palermo	Sebina	P	P
16	41,5	Polo SBN Università di Roma "La Sapienza". 2	Sebina	R	P R U Z
17	41,4	Provincia di Trapani	Sebina	P	P
18	41,1	Libera università di lingue e comunicazione IULM	ALEPH	N	U
19	40,8	Provincia di Brindisi	Sebina	P	E P
20	40,3	Provincia di Venezia	Easyweb	P	P
21		Università di Genova	ALEPH	R	U

\* Copertura: N – nazionale; P – provinciale; R – regionale.  
 \*\* Biblioteche: E – ecclesiastiche; N – statali; P – pubbliche; R – di enti di ricerca; S – scolastiche; U – di università; Z – altre

l'essere sfruttata in modo ideale. Nella classifica, la seconda posizione assoluta è occupata a pari merito dai due OPAC che usano il software dell'Università di Firenze, che ha quindi il miglior indice medio: 51,8. I due OPAC considerati sono quelli dei poli SBN universitari di Bologna e di Padova, mentre quello dell'Università di Firenze stessa non è stato incluso nell'indagine perché riferito a biblioteche di una sola città (vedi box a pagina 24). Tutte le altre posizioni dalla prima alla diciassettesima sono invece occupate da OPAC Sebina, un software che si dimostra quindi più attento agli aspetti semantici rispetto ai concorrenti: ha una media complessiva di 38,3, contro 33,4 di Aleph, 30,4 di Webif, 28,0 di Bibliowin, 26,5 di Easyweb. La qualità dell'indicizzazione semantica dipende comunque dalle particolari configurazioni adottate da ciascun catalogo: infatti un particolare OPAC Sebina, quello del polo SBN di Romagna, si eleva nettamente su tutti gli altri raggiungendo l'indice migliore in assoluto con 54,7. Tra i software meno diffusi si distinguono per i loro buoni valori Book-

markweb (32,8), Zetesis (32,4) e Librinlinea (30,8). Aldilà delle singole realizzazioni, è interessante osservare quanto emerge raggruppando gli OPAC in base ad alcune loro caratteristiche registrate nel repertorio AIB-CILEA degli OPAC italiani.<sup>10</sup> Ad esempio, il gruppo dei cataloghi con indice di semanticità medio più alto (31,9) è quello degli OPAC a copertura regionale, ossia comprendenti biblioteche di più comuni di province diverse, mentre quelli comunali e quelli nazionali hanno valori molto inferiori. La ripartizione per regione dà i valori massimi a Campania e Calabria, grazie all'unico OPAC esaminato in entrambi i casi, mentre fra le regioni con più cataloghi compresi nel campione si piazzano bene Marche, Lazio, Sicilia ed Emilia-Romagna e male (anche per la diffusione degli Easyweb) la Toscana. Il raggruppamento per tipo di biblioteche assegna la media migliore a quelle statali (33,1), seguite dalle pubbliche (30,6). Va però considerato che ciascun OPAC può comprendere biblioteche di diverse tipologie: se si isolano gli OPAC "puri"

con biblioteche di una sola tipologia i risultati cambiano, vedendo ora prevalere le pubbliche (30,0), seguite dalle scolastiche (24,7) e scendere molto le statali (15,5); in entrambi i casi vanno abbastanza male le biblioteche degli enti di ricerca diversi dalle università.

### Per non concludere

L'indagine è giunta dunque a raccogliere una serie di dati interessanti e soprattutto – si spera – utili a riflettere sul funzionamento degli OPAC attuali e a progettare quelli futuri. In questo senso non si rivolge solo ai bibliotecari in senso stretto, ma vorrebbe offrirsi come uno strumento alla collaborazione fra bibliotecari e informatici sul delicato fronte delle informazioni fornite ai lettori. È stata una scelta coerente con questi scopi la pubblicazione in rete degli utensili della ricerca: lista di controllo, foglio di calcolo, formula dell'indice di semanticità. In questo modo, chiunque lo desideri potrà saggiare il funzionamento degli accessi semantici in OPAC non esaminati in questa pri-

ma fase: sarà naturalmente gradito l'invio dei dati relativi perché siano integrati con quelli già raccolti. Un'altra prospettiva interessante è quella di sottoporre il medesimo campione di OPAC a una nuova analisi, fra qualche tempo, per verificare gli eventuali cambiamenti. Il sito web rimane, ed è aperto a futuri arricchimenti, come punto di riferimento per gli interessati, come raccolta di documentazione e di citazioni bibliografiche, come opera in continuazione. E come stimolo a realizzare cataloghi con indici di semanticità sempre migliori.

#### Note

<sup>1</sup> TIM BERNERS-LEE – JAMES HENDLER – ORA LASSILA, *The semantic Web*, "Scientific American", (May 2001), 284, p. 34-43.

<sup>2</sup> *OPAC semantici: indagine sugli accessi semantici nei cataloghi in rete italiani*, Università di Pavia, 2003, <<http://mate.unipv.it/biblio/sem/>>.

<sup>3</sup> HENRI VAN DE WAAL, *Iconclass*, ed. italiana a cura di Marco Lattanzi, Simona Ciofetta, Elena Plances, Roma, Istituto centrale per il catalogo e la documentazione, 2000, 8 v.

<sup>4</sup> "While DDC, LCC and LCSH are common in libraries, alphabetical entry terms to the classification scheme or subject headings are often lacking and users have to fall back on 'keywords' – a term that is a particular bête noire for me because I never know whether it means controlled descriptors or just words taken from the titles or text", LEONARD WILL, *Facets of the future: subject organisation in the Internet age*, "The Bliss classification bulletin", 45 (2003), p. 11-18 (11).

<sup>5</sup> ANDREA FABBRIZZI, *L'applicazione delle norme GRIS in CDS-ISIS TECA*, in *L'indicizzazione per soggetto della sezione locale: una applicazione delle norme GRIS*, a cura di Massimo Fedi e Raffaella Marconi, con la collaborazione di Andrea Fabbrizzi, Marta Gori, Paolo Panizza, Firenze, Provincia di Firenze, Biblioteca comunale di Bagno a Ripoli, Biblioteca comunale di Fiesole, 2000, p. [89]-109.

<sup>6</sup> MARIA CHIARA GIUNTI, *In SBN con Dewey: il catalogo classificato del Polo della Biblioteca nazionale centrale di Firenze*, "Bollettino AIB", 41 (2001), 1, p. 31-46.

<sup>7</sup> *I diritti di Dewey*, "AIB notizie", 4 (1992), 10, p. 10.

<sup>8</sup> ASSOCIAZIONE ITALIANA BIBLIOTECHE. GRUPPO DI RICERCA SULL'INDICIZZAZIONE PER SOGGETTO, *Guida all'indicizzazione per soggetto*, Roma, Associazione italiana biblioteche, 1996 (rist. con correzioni, 2001), p. 59; CLAUDIO GNOLI, *Coordinazione, ordine di citazione e livelli integrativi in ambiente digitale*, "Bibliotime", n.s., 6 (2003), 1, <<http://www.spbo.unibo.it/bibliotime/num-vi-1/gnoli.htm>>.

<sup>9</sup> 5% per le informazioni e altrettanto per gli strumenti ausiliari; 20% ciascuno agli accessi per soggetti e per classi; per le visualizzazioni disponibili: 25% a quella in liste, 5% in schede sintetiche, 10% in schede dettagliate; infine 10% alla navigabilità delle schede.

<sup>10</sup> *OPAC italiani*, a cura di Antonella De Robbio, Claudio Gnoli, Paola Rossi e altri, "AIB-Web", 1999, <<http://www.aib.it/aib/opac/opac1.htm>>.