

Here's the hard challenge. Figure out what has in fact changed, and what is changing rapidly. Understand that arguments like 'not everything is digitized' and 'not everyone has computers', though true, are less effective arguments with each passing year.

And then, in that light, built the library back up. Figure out what it does that won't be washed away over time. Figure out what can be jettisoned now, to save funds and effort for more effective opportunities. Figure out what libraries aren't doing, or aren't doing very much of, that can provide new value and new reasons for existing. In short, look into the abyss and start building a bridge over it» (Tim Spalding, in <http://www.librarything.com/topic/93959>, message 20).

Anna Galluzzi

Biblioteca del Senato "Giovanni Spadolini"

Roberto Raieli. *Nuovi metodi di gestione dei documenti multimediali: principi e pratica del MultiMedia Information Retrieval*, prefazione di Giovanni Solimine. Milano: Editrice Bibliografica, 2010. 390 p. ISBN 978-88-7075-690-6. 28,00.

Per millenni le biblioteche hanno conservato documenti prevalentemente testuali, sebbene talvolta arricchiti da illustrazioni o, addirittura, a carattere principalmente iconografico (dagli atlanti ai fumetti) o, più recentemente, anche sonoro (nastri e dischi in vari formati) o video (su pellicola, VHS o altri supporti). Per permettere agli utenti di individuare quelli di volta in volta desiderati sono stati utilizzati metadati (cioè rappresentazioni sintetiche, strutturate e normalizzate dei documenti stessi) e indici (cioè insiemi organizzati e interrogabili di metadati) entrambi testuali, ovvero formati da stringhe più o meno estese e complesse di lettere dell'alfabeto, di numeri e di segni di punteggiatura, "estratti" dai documenti stessi o "assegnati" ad essi dai bibliotecari. Tale procedura è sempre apparsa a tutti quella al tempo stesso più naturale (vista l'omogeneità "materiale" fra i macrotesti indicizzati e i microtesti utilizzati per indicizzarli) e più facilmente realizzabile (viste le tecnologie pre-elettroniche disponibili), ed è quindi stata quella di gran lunga prevalente, con sporadiche esperienze di classificazioni tramite simboli o colori, rivolte prevalentemente all'infanzia.

Dopo l'invenzione del computer, con l'invasione dei documenti digitali (sia quelli ricavati da originali analogici che quelli "digitali nativi", entrambi resi disponibili *online* oppure memorizzati su supporti locali come Cd e Dvd) le biblioteche hanno visto rapidamente incrementare, nelle collezioni da loro possedute o comunque gestite, sia il numero (in termini assoluti) che la percentuale (in termini relativi) di documenti grafici, audio, video e più propriamente multimediali, ovvero composti, in varie proporzioni e abbinamenti, sia da parti testuali che da suoni e da immagini sia statiche e in movimento.

Anche tali documenti multimediali possono essere corredati da metadati e indici testuali, scorribili o interrogabili – sia in ambiente analogico che digitale – in modo sostanzialmente tradizionale, ma la radicale eterogeneità "materiale" che, stavolta, si verifica fra i microtesti utilizzati per l'indicizzazione e i non-testi indicizzati, rende questo approccio meno proficuo rispetto a quanto avviene abitualmente con i documenti testuali, come evidenziato ad esempio dall'impossibilità di "estrarre" parole chiave da suoni e immagini o dalla problematicità di applicare il concetto di "soggetto" a una pittura astratta o a una sinfonia.

Per fortuna il problema ha portato con sé anche la propria soluzione, grazie alla possibilità tecnologica (limitatamente ai documenti multimediali di tipo digitale) di metodi di ricerca documentaria radicalmente diversi da quelli tradizionali, resi possibili dall'omogeneità del formato digitale sia dei documenti ricercati che dei metadati utilizzati nella ricerca stessa. Grazie alla codifica binaria l'intrinseca pluralità tipica dei contenuti multimediali in ambiente tradizionale (un bassorilievo in bronzo e un disco in vinile sono oggetti ben diver-

si fra loro) viene ridotta ad unità (una immagine digitale e un suono digitale possono essere ospitati nel medesimo Dvd), secondo il noto fenomeno della “convergenza al digitale”, che rende possibile il *multimedia information retrieval* (MMIR), ovvero la ricerca di documenti non testuali con tecniche basate prevalentemente sul loro stesso contenuto multimediale, scavalcando l’intermediazione dei metadati testuali. In sostanza, alle classiche interfacce delle banche dati testuali, interrogabili coi metodi dell’*information retrieval* (IR) tradizionale attraverso indici composti esclusivamente da parole, è ora possibile aggiungere interfacce che permettano di effettuare ricerche in diverse “dimensioni”, non solo tramite testi ma anche attraverso immagini e suoni. Nell’ambito di questo approccio *content-based* – complementare rispetto a quello classico, *term-based* – aumenta il rilievo delle tecniche di scorrimento e navigazione (tipiche dell’ambiente analogico) rispetto a quelle di vera e propria interrogazione, nate in ambiente digitale. E comunque le interrogazioni, quando presenti, non potranno basarsi, come sarebbe usuale, sull’esatta corrispondenza fra il termine ricercato e quello recuperato, ma dovranno spesso muoversi entro un certo arco di similitudine.

Il libro di Roberto Raieli, bibliotecario che di MMIR studia e scrive da tempo e che ha recentemente anche conseguito un dottorato su tali temi, è un’ottima introduzione ai concetti, le tecniche, le realizzazioni, i problemi, la storia e le prospettive del MMIR, che vengono illustrati dettagliatamente ma senza mai perdere accessibilità e chiarezza. Il volume è strutturato in due parti: la prima (di quasi 200 pagine) dedicata agli aspetti più generali e teorici e la seconda (di quasi 140 pagine) che offre invece una carrellata sulle principali applicazioni pratiche già realizzate o in via di sviluppo negli Stati Uniti ed in Europa, con un doveroso occhio di riguardo per l’Italia ed includendo anche biblioteche digitali non convenzionali (legate ad esempio ai settori dell’industria tessile e della moda, dello spettacolo e dell’arte, delle cure ospedaliere e delle indagini giudiziarie) e i motori per ricerca delle immagini disponibili sul web.

Molto efficace e documentata risulta l’analisi sia dei limiti dell’IR classico applicato ai documenti multimediali che delle innovative opportunità offerte dal MMIR, articolato nelle sue varie componenti: *audio retrieval*, *visual retrieval* e *video retrieval*, ma anche *text retrieval*, in quanto «il diffuso metodo dell’interrogazione free-text sul testo pieno di un documento è già una modalità di ricerca content-based, in quanto fondata sul trattamento diretto dei dati di contenuto. Infatti, il content-based retrieval di dati testuali è un metodo di ricerca che usa il testo effettivo dei documenti piuttosto che un qualunque metadato descrittivo ‘aggiunto’ a esso, come, appunto, la ricerca free-text, che è indipendente da qualunque vocabolario e scandaglia direttamente il full text dei documenti» (p. 347).

Particolarmente rappresentativo delle novità introdotte dal MMIR è il metodo di ricerca con cui gli utenti propongono al sistema informativo (scegliendolo fra quelli disponibili o aggiungendolo essi stessi, ad esempio disegnando o canticchiando) una entità dotata di caratteristiche multimediali (colore, forma, timbro, melodia, ecc.) simili a quelle dell’entità ricercata, senza bisogno di dover indovinare come tali caratteristiche potrebbero essere state denominate dai relativi indicatori. Il sistema, tipicamente, risponde agli utenti proponendo loro sia un insieme di entità con caratteristiche simili al campione iniziale che una interfaccia attraverso cui gli utenti potranno eventualmente proseguire ulteriormente la ricerca modificando i parametri utilizzati dagli algoritmi del sistema per tradurre in termini matematici il campione.

Raieli, pur sottolineando i numerosi ambiti in cui l’introduzione del MMIR potrebbe risultare fin d’ora particolarmente vantaggiosa sia per gli utenti che per i gestori dei sistemi documentari multimediali, non nasconde i vari aspetti ancora non completamente convincenti delle attuali realizzazioni, che vanno dalla scarsa usabilità di molte delle interfacce, non sufficientemente intuitive, all’eccesso di rumore o – inversamente – di perdita dell’informazione, di cui possono essere responsabili alcuni degli algoritmi di ricerca. Il principale limite, però, del MMIR, ovvero il “gap semantico” consistente nella «distanza tra la rappresentazione di ‘alto livello’ concettuale di un oggetto propria della conoscenza umana, e la denota-

zione di 'basso livello' formale di cui è capace autonomamente la macchina» (p. 138), non è superabile da nessuna futuribile meraviglia ingegneristica ma può essere evitato solo da una organica integrazione fra ricerche *content-based* e *term-based*, magari grazie ad interfacce che permettano di incrociare ogni genere di interrogazione.

Particolarmente apprezzabile, nella trattazione di Raieli, è l'equilibrio fra innovazione e tradizione, evidente non solo nel modo in cui il "gap semantico" viene riconosciuto e risolto, ma anche nell'ammissione della sostanziale continuità fra IR e MMIR, visto che in entrambi i casi una ricerca davvero efficace è possibile solo se preventivamente sono stati creati gli opportuni metadati e i relativi indici interrogabili, comunque necessari anche se estratti automaticamente e non composti da testi ma da suoni e immagini particolarmente "concisi" e rappresentativi. Continuità riscontrabile, peraltro, anche nel ricordare, in vari punti del libro, come anche in ambito multimediale i metadati aggiunti estemporaneamente dagli utenti, oggi giorno così di moda, possano assumere una utile funzione integrativa, ma giammai sostituire *in toto* quelli prodotti tramite sistematiche procedure professionali.

I punti deboli del volume, pur complessivamente utile e raccomandabile per chiunque voglia tenersi aggiornato sulle nuove frontiere della ricerca documentaria, sono il titolo, che promette più di quanto il libro stesso mantenga, e il primo capitolo, intitolato *Un contesto culturale e scientifico per la teoria del MMIR*, a mio avviso non particolarmente organico rispetto al resto della trattazione, in cui vengono esaminate una decina di pubblicazioni italiane (di Balsamo, Bisogno, Castellucci, Miccoli, Salarelli, Serrai e Solimine) relative ai contenuti, ai confini e ai reciproci rapporti delle discipline della biblioteconomia, della documentazione e delle scienze dell'informazione. I testi sono tutti importanti, ben scelti, ben riassunti e ben commentati, ma il nesso con le tematiche a cui è dedicato il resto del libro non è a mio avviso particolarmente forte nè, peraltro, lo stesso Raieli insiste più di tanto per riscontrarlo, al di là della pacifica conclusione che il MMIR fa ovviamente parte, a buon diritto, dell'ambito disciplinare così attentamente analizzato. Per quanto riguarda il titolo complessivo del volume (*Nuovi metodi di gestione dei documenti multimediali*), l'inclusione del termine "gestione" avrebbe potuto far pensare a sostanziosi capitoli dedicati anche alla selezione, all'acquisizione, alla conservazione e alla fruizione dei documenti stessi, tutti temi invece sostanzialmente non affrontati, anche perchè le 11 pagine del conclusivo paragrafo 7.3, intitolato *I rapporti del MMIR con gli altri sistemi di gestione documentale*, si occupano (correttamente, visto l'impianto complessivo del volume) dell'interoperabilità fra i metadati gestionali contenuti negli OPAC e in altri tipi di indici, piuttosto che di gestione complessiva dei documenti. La bibliografia, ben selezionata, è utilmente suddivisa in sei sezioni tematiche (MMIR, *text retrieval*, *visual retrieval*, *video retrieval*, *audio retrieval*, IR e argomenti correlati). Vista la grande quantità di termini tecnici sarebbe stato particolarmente utile un indice analitico.

Riccardo Ridi
Università Ca' Foscari di Venezia

Les collections électroniques, une nouvelle politique documentaire, sous la direction de Pierre Carbone, François Cavalier. Paris: Editions du Cercle de la Librairie, 2009, 311 p. (Bibliothèques). ISBN 978-2-7654-0975-5. 39,00.

I testi digitali, i loro autori, le biblioteche, gli editori, la rete. Gli autori del libro curato da Pierre Carbone e François Cavalier – ventisei, tutti professionisti dell'informazione – trasferiscono questi temi nell'ambito della gestione delle collezioni e ne valutano l'impatto sulla professione bibliotecaria, sulle istituzioni culturali e sulle modalità di apprendimento e di ricerca dei loro utenti.

Il libro si può leggere secondo l'ordine scelto dai curatori (4 parti, 17 capitoli) oppure seguendo lo sviluppo di alcuni temi (*l'open access*, le licenze, i consorzi). In ogni caso, *Les*