

# ***Funzionalità e flessibilità di schemi di classificazione tradizionali applicati a un CMS (Content Management System): faccette, CDD, JITA***

**Susanna Dal Porto<sup>1</sup> e Andrea Marchitelli<sup>2</sup>**

Relazione presentata al 2. seminario ISKO Italia-UniMIB, 9 giugno 2006, Milano  
<<http://www.iskoi.org/doc/milano06.htm>>

## **INTRODUZIONE**

L'impiego di schemi di classificazione bibliografica nel *web* si è fatto via via più ricorrente; in particolare tali schemi, con i necessari adattamenti, sono impiegati nel *web* per l'organizzazione della conoscenza, cioè per dare una struttura coerente e organica ai contenuti presenti nel web. Molti miglioramenti devono ancora essere apportati a questi schemi affinché diventino degli strumenti validi e flessibili per la gestione della conoscenza.

Questo intervento si pone come obiettivo quello di illustrare tre schemi di classificazione diversi così come sono stati applicati a tre *weblog* diversi, evidenziandone le funzionalità e le caratteristiche nonché i limiti e spiegando per quale motivo sia stato scelto uno schema piuttosto che un altro.

I tre *weblog* presi in considerazione hanno un layout simile a tutti i *blog*: nella home page sono visibili una serie di *post* ordinati dal più al meno recente. Un meccanismo di archiviazione fa sì che dopo una certa permanenza sulla home page i post siano trasferiti in pagine di archivio. Ogni post ha una data di pubblicazione, un titolo e un *permalink*, cioè un proprio URL specifico al quale è sempre possibile reperire il documento. Inoltre un *weblog* presenta varie possibilità di ricerca dei post; le più comuni

---

<sup>1</sup> Università di Pisa, Dipartimento VI - Sistema Bibliotecario, Archivistico e Museale - <http://biblio.unipi.it/>  
LUNGARNO PACINOTTI, 43/44 - 56126 PISA  
[susanna.dalporto@jus.unipi.it](mailto:susanna.dalporto@jus.unipi.it)

<sup>2</sup> CILEA - <http://www.cilea.it> - Piazza G. Marconi, 10, I-00100 ROMA  
Unità Operativa per la realizzazione di servizi a supporto delle biblioteche e della loro utenza  
[marchitelli@cilea.it](mailto:marchitelli@cilea.it)

sono: ricerca nel full text dei post, browse per mese di pubblicazione o per titolo. Si tratta dunque di una struttura assai semplice ed uniforme, che rimane costante: a variare saranno i particolari contenuti di ogni articolo e la sua lunghezza.

Si può osservare, ad esempio, l'architettura dell'informazione di Biblio(a)tipici, che si presenta con una home page strutturata in quattro zone:

- la testata (*header*)
- le due colonne centrali
- il pie' di pagina (*footer*)

L' *header* contiene gli elementi identificativi del blog: il titolo e una breve tagline. Inoltre, tutti gli elementi che facilitano l'accesso ai post, in particolare la barra di navigazione, sono concentrati in questo spazio di massima visibilità.

Sotto l'*header*, la colonna di sinistra, larga per i due terzi dello schermo, contiene gli ultimi 7 post, disposti in ordine cronologico inverso. Ogni post è caratterizzato, oltre che da un permalink (URI stabile, identificante i singoli articoli), dal link alla form per aggiungere commenti, da quello alla versione per la stampa e dall'indicazione delle categorie secondo le quali il post è classificato (ma su questo torneremo più oltre).

La colonna di destra contiene alcuni servizi legati alla vita del blog: la form di ricerca semplice, il calendario del mese corrente, che permette di accedere ai post di una certa giornata, l'elenco degli ultimi commenti, che contribuisce a sviluppare una dimensione comunitaria, l'elenco di concorsi e altre offerte di lavoro in scadenza e, da pochi giorni, il link della settimana, una risorsa selezionata per il repertorio e messa in evidenza per qualche giorno.

Il *footer* , infine, oltre al link al blog di Andrea Marchitelli, ideatore di Biblio(a)tipici, indica gli standard di riferimento (XHTML e CSS 2) e l'adesione del gruppo a CABI<sup>3</sup>.

---

<sup>3</sup> La Biblioteca Nazionale Marciana di Venezia ha avviato il Progetto CABI, iniziato con la Campagna per l'Accessibilità delle Biblioteche in rete <<http://marciana.venezia.sbn.it/CABI/>>, una campagna rivolta alle biblioteche, ma che vuole interessare tutti i "fornitori" di Cultura e, naturalmente, anche tutta la Pubblica Amministrazione per informare e sensibilizzare sul problema dell'accessibilità

Per quanto riguarda l'aspetto tecnologico, i tre *weblog* che vedremo sono realizzati con un CMS (*Content Management System*).

Usabilità e accessibilità sono state, fin dall'inizio, obiettivi da raggiungere: l'uso di un CMS ha permesso di coinvolgere nella redazione persone che non avevano una conoscenza approfondita dell'HTML, mantenendo comunque un buono standard di strutturazione e compilazione delle pagine. Un CMS si occupa, infatti, di tutto il processo gestionale dei contenuti di un sito complesso (un blog in questo caso), dalla raccolta di contenuti creati o recuperati da fonti esterne, alla gestione tramite basi di dati, alla loro archiviazione (*storage*) ed infine alla loro pubblicazione. La pubblicazione costituisce un aspetto centrale di un CMS: essa si basa sulla creazione iniziale di *templates*, cioè di modelli che definiscono l'impaginazione dei testi, recuperandone il contenuto dal database.

Dunque questo aspetto dei CMS di separare la creazione e la gestione dei contenuti dalla loro pubblicazione, tramite la definizione di modelli di pubblicazione che possono essere pre-stabiliti, e possono essere modificati, costituisce una grande opportunità per i *content managers* perché li rende completamente autonomi.

Le pagine web dei blog sono redatte in XHTML, con una separazione completa del contenuto dalla presentazione grafica, attraverso l'uso di due diversi fogli di stile, uno dei quali utilizzato per la costruzione di una versione stampabile dei singoli post.

Il codice viene periodicamente verificato con il validatore W3C, accessibile a <<http://validator.w3.org>>. In occasione dei diversi restyling, l'accessibilità del sito è stata testata anche con un software apposito, che permette di simulare la resa grafica dei diversi browser.

Il CMS utilizzato per i tre blog è *MovableType* <<http://www.sixapart.com/movabletype/>>, nella versione 3.2, che è dotato di un'interfaccia web intuitiva e di molte funzionalità innovative come il meccanismo di *trackback* che consente di tenere traccia delle citazioni da un blog all'altro: se il post di un blog cita un post di un altro blog, quest'ultimo blog viene informato della citazione. Un esempio di citazione 'interna' cioè che rimanda a due post dello stesso blog si può vedere qui: [http://www.biblioatipici.it/2005/11/nuova\\_lista\\_di.shtml](http://www.biblioatipici.it/2005/11/nuova_lista_di.shtml)

Infine, una caratteristica fondamentale di MovableType è che esso prevede, così come altri CMS, un modulo apposito per la gestione di **categorie** su n livelli, cioè gerarchicamente organizzate, con la possibilità di selezionare più categorie per ciascuna *entry* (o post, nel nostro caso). Questa funzionalità era già prevista nelle versioni precedenti del software mentre la possibilità di creare con facilità un numero infinito di sotto-categorie è stata implementata solo a partire dalla versione 3.1. Da allora, il menu a tendina utilizzato per assegnare a ciascuna *entry* una o più categorie non è stato più capace di gestire queste potenti opzioni e, nella versione 3.2, si è optato per un menu ad albero.

Le funzionalità appena descritte, offerte anche da altri CMS, sono generalmente utilizzate nei blog per assegnare a ciascun post, a seconda del suo contenuto, una o più categorie o sottocategorie, così da consentire un raggruppamento dei post per argomento e una chiave di accesso semantica all'archivio dei post.

Nei blog che andremo a vedere si è deciso di applicare a questa struttura categoriale gerarchica, propria di tutti i CMS, tre schemi di classificazione bibliografica sviluppati e utilizzati nel mondo bibliotecario.

L'impiego sperimentale di schemi di classificazione bibliografica nel web si è fatto via via più frequente considerato che la teoria classificatoria semantica ha una lunga storia all'interno della comunità LIS (Library and Information Science) e le classificazioni sviluppate nella lunga storia della biblioteconomia, pur rinunciando all'assoluta scientificità, mostrano un livello di *coerenza* maggiore rispetto agli schemi *home-grown* del Web. Uno schema di classificazione coerente rispetta la regola dell'*unico fundamentum divisionis*, cioè ogni nodo della classificazione produce categorie mutuamente esclusive. Inoltre la struttura degli schemi di classificazione che vedremo si sposa assai bene con qualunque sistema basato su database.

La scelta di applicare tre schemi di classificazione diversi alla medesima struttura dei contenuti consente di evidenziare i vantaggi dell'impiego di uno schema piuttosto che di un altro.

L'ordine di presentazione dei blog sarà il seguente: dallo schema di classificazione più complesso ma dotato di maggiori potenzialità, cioè lo schema basato

sul principio delle faccette, a quello più semplice ma meno funzionale (JITA), passando attraverso uno schema tradizionale di classificazione bibliografica: la Classificazione Decimale Dewey.

## 1. SCHEMA DI CLASSIFICAZIONE A FACLETTE

Biblioatipici <<http://www.biblioatipici.it/>> è un weblog dedicato al mondo dei lavoratori 'atipici' in biblioteche e centri di documentazione italiani. È mantenuto da alcuni bibliotecari per altri bibliotecari o aspiranti tali allo scopo di condividere conoscenza su temi comuni attinenti al lavoro e alla professione. È nato nell'aprile del 2004 e, ad oggi, ha un bacino di utenti di circa 300 contatti giornalieri.

Biblioatipici è un weblog collaborativo. Dal punto di vista dei contenuti non è facile classificarlo perché è un weblog di rassegna ma anche di segnalazione, molto di commento, un poco anche di narrazione e infine di progetto.

Attualmente i post presenti nel blog sono quasi 500, dunque un numero che comincia ad essere consistente e richiede tecniche di *retrieval*.

Nella home page è disponibile la funzionalità 'search' che cerca nel *full text* dei post. Inoltre i post, come si può vedere dalla pagina dell'archivio <<http://www.biblioatipici.it/archives.shtml>>, sono **scorribili** per:

- popolarità (vi si trovano i 20 post più recenti con il loro indice di popolarità, legato alla quantità di commenti ricevuti);
- categoria;
- mese di pubblicazione;
- elenco alfabetico per titoli di tutti i post

### Le categorie

Lo scorrimento per categorie è sicuramente la chiave di ricerca più flessibile e funzionale.

Un progetto iniziato nell'estate del 2005 e tuttora in corso di perfezionamento è legato all'applicazione alle categorie del CMS di uno schema di classificazione basato sul principio delle faccette.

La scelta di utilizzare uno schema di classificazione di questo tipo piuttosto che uno di tipo gerarchico-enumerativo, ha diverse motivazioni: in primo luogo, in questo blog si indicizzano contenuti specializzati e le faccette funzionano particolarmente bene per un insieme omogeneo di oggetti; in secondo luogo, poiché Biblioatipici è un blog collaborativo, lo schema a faccette consente a coloro che indicizzano di fornire un accesso multiplo ai vari post, cioè un approccio semantico multidimensionale, soddisfacendo così modelli mentali e strategie di accesso anche molto diverse tra loro. Infine, questo blog è destinato a crescere e modificarsi continuamente con una possibile revisione o integrazione a posteriori dei criteri di classificazione; pertanto uno schema a faccette è la miglior opzione possibile perché è flessibile e scalabile.

Nell'applicare lo schema a faccette al blog si è utilizzata la seguente **metodologia**: nella prima fase, a livello del piano concettuale, il contenuto semantico di ciascun post è stato scomposto in parti semplici, detti *isolati*; poi, ripetendo la stessa analisi per più post si è visto che tutti gli isolati tendevano a raggrupparsi in un numero limitato di *categorie* affini che quindi sono state poi identificate per via induttiva.

A questo proposito occorre sottolineare che è stato possibile applicare uno schema di classificazione a faccette al blog Biblioatipici, proprio perché quando si è cominciato ad applicarlo, il blog aveva già una certa consistenza di post e quindi le categorie e le faccette sono emerse proprio dalla letteratura stessa, quindi dal basso, dalla pratica del classificare. Diversamente, l'applicazione di uno schema di classificazione gerarchico-enumerativo come la CDD sarebbe stata possibile anche con un solo post in archivio (classificazione che viene dall'alto).

Le *categorie fondamentali* individuate in Biblioatipici sono sette: numericamente più di quelle di Ranganathan e meno di quelle elaborate dal CRG (Classification Research Group), che sono tredici. Le categorie fondamentali di Ranganathan sono espresse dalla formula **PMEST** (Personalità, Materia, Energia, Spazio, Tempo) e furono riviste alla luce degli studi del CRG. Nell'ambito del CRG,

infatti, il lavoro si concentrò sull'identificazione di categorie più specifiche rispetto a quelle di Ranghanathan, così da facilitare l'esercizio dell'analisi categoriale. D'altra parte la gamma completa di categorie dell'elaborazione del CRG è usata probabilmente soltanto nei soggetti tecnologici o comunque particolarmente complessi.

Nel blog Biblioatipici lo schema è stato semplificato, trattandosi di un sito che accoglie informazioni, materiali, segnalazioni già fortemente selezionate.

Le categorie standard individuate sono le seguenti <<http://www.biblioatipici.it/archives.shtml#cat>>:

1. **Entità** = questa categoria è presente anche nello schema del CRG ed è la prima anche in quello schema. È equivalente all'uso più semplice (perché ne costituisce solo una parte) della categoria 'Personalità' di Ranganathan e riguarda l'interesse principale o l'oggetto di qualsiasi disciplina (le piante in botanica, le sostanze in chimica, le nazioni in storia). Questa categoria contiene perlopiù oggetti fisici, oppure aggregazioni di oggetti in sistemi.

La categoria **Materiale** presente nello schema del CRG ed equivalente alla categoria 'M=Materia' della Classificazione Colon non è stata inclusa in questo schema perché scarsamente significativa in questo contesto. Un esempio: se l'Entità è casa, i Materiali sono: legno, mattoni, ecc..

2. **Attività** = corrispondente all' 'Energia' di Ranghanathan, ha due equivalenti nello schema del CRG: 'Processi' e 'Operazioni'. Il CRG, infatti, distingue tra il 'Processo' che è rappresentato da azioni intrinseche e spontanee e le 'Operazioni', cioè azioni determinate da un agente esterno. Nel nostro schema tale distinzione non è stata mantenuta perché non è significativa: tutti i concetti in questa categoria sono 'Operazioni'. Si possono trovare i 'Processi' soprattutto nelle scienze fisiche e naturali.
3. **Prodotti** = (presente nello schema del CRG) esiti o risultati di processi in, o di operazioni su, entità; solitamente consistono in prodotti fisici. Anche se questa categoria è in gran parte limitata all'area della tecnologia, ed è generalmente assente nelle arti e nelle discipline umanistiche e sociali, trova impiego in questo blog perché ci sono molto oggetti che costituiscono il risultato dell'attività del gruppo.

4. **Strumenti** = vd. la seguente
5. **Agenti** = i mezzi attraverso i quali delle operazioni vengono effettuate. Gli Agenti possono essere distinti in 'Persone' e 'Strumenti' e, a un livello complesso, possono essere rappresentati da istituzioni. Le due categorie di agenti possono presentarsi insieme; es. in medicina un chirurgo (Agente persona) può asportare del tessuto utilizzando un laser (Agente strumento). In questo schema gli agenti 'Persona' e 'Strumento' vengono separati a differenza dello schema del CRG che prevede una sola categoria.
6. **Spazio** = qualsiasi tipo di dimensione politica, fisiografica o spaziale (es.: gli USA, montuoso, interno). In questo blog questa categoria è molto utile perché, per esempio, è possibile riunire tutti i concorsi secondo l'area geografica.
7. **Tempo** = qualsiasi tipo di caratteristica storica, cronologica o temporale (es.: medievale, permanente, notturno). Questa categoria è stata poco o per nulla sviluppata (contiene un solo post) perché il blog Biblioatipici ha ancora una storia piuttosto recente e anche perché tra le possibilità di ricerca dei post c'è anche quella di cercarli nell'*archivio cronologico*, dove i post si trovano suddivisi per mesi a seconda della data di pubblicazione. Ciò non vuol dire che un post pubblicato a marzo ad esempio riguardi la segnalazione di un concorso che si svolge a marzo, ma, con una certa approssimazione e scarto, resta comunque una buona chiave di ricerca temporale, per il momento.

Ciascuna categoria contiene al suo interno le **facette** che sono state individuate. Ad esempio la categoria 'Agenti' ha come faccette 'Sindacato', 'Enti privati', 'Enti pubblici', 'Associazioni professionali', 'Atipici'. Risulta evidente che le faccette sono **mutuamente esclusive** l'una rispetto alle altre, infatti non vi è sovrapposizione semantica.

All'interno di alcune di queste faccette sono state individuate delle **facette di secondo livello**. Ad esempio: la faccetta 'Associazioni professionali' ha le seguenti sottoclassi: 'IAML-Italia', 'AIDA', 'AIB'. In questi casi la faccetta di primo livello funge da **classe di raggruppamento** e quindi non contiene post o ne contiene molto pochi perché, nel classificare un post, si cerca sempre di utilizzare la classe più specifica.

Quindi ciascun post è stato classificato con una o più faccette a seconda della complessità dell'argomento. Si veda ad esempio questo post:

<[http://www.biblioatipici.it/2006/05/concorso\\_alluni\\_18.shtml](http://www.biblioatipici.it/2006/05/concorso_alluni_18.shtml)> per il quale è stato possibile individuare le seguenti categorie: 'Attività', 'Prodotti', 'Agenti', 'Spazio'; quindi nella stringa di classificazione avremo le seguenti faccette: H11: Concorsi non scaduti::J94: Tempo indeterminato::M58: Università::S51: Italia settentrionale.

Nella stringa di classificazione di ciascun post deve essere rispettato un **ordine di citazione** cioè un ordine di precedenza tra le faccette. Questo ordine rispetta la regola di Ranganathan della "*concretezza decrescente*": ciò significa che le faccette che compaiono per ultime nell'ordine di citazione sono quelle considerate meno rilevanti per la specificazione del soggetto. L'ordine di precedenza tra le faccette del blog Biblioatipici rispetta, quindi, l'**ordine di citazione standard** che colloca in questa successione le *categorie fondamentali*: 'Entità', 'Attività', 'Prodotti', 'Strumenti', 'Agenti', 'Spazio', 'Tempo'. Per fare un esempio il seguente post: <[http://www.biblioatipici.it/2006/03/nuovo\\_disegno\\_d.shtml](http://www.biblioatipici.it/2006/03/nuovo_disegno_d.shtml)> dal punto di vista semantico è scomponibile in quattro *isolati* afferenti alle seguenti categorie: 'Entità', 'Agenti', 'Spazio'; di conseguenza nella stringa di classificazione l'ordine delle faccette sarà il seguente: B6: Bibliotecari::M1: Atipici::M59: Enti locali::S58: Isole.

Come fase intermedia nell'applicazione dello schema, si è individuata una notazione espressiva capace di produrre un ordine significativo utile per il browsing. Infatti, affinché il sistema restituisca automaticamente l'ordine di citazione standard nelle stringhe di classificazione, ci si è serviti di una notazione mista di lettere e numeri, che permette la strutturazione gerarchica delle classi e l'applicazione di un criterio organico di ordinamento. Le categorie quindi sono individuate con una lettera e le faccette di primo e secondo livello con lettere e numeri.

Alle categorie (e alle faccette di conseguenza) sono state assegnate delle lettere in ordine alfabetico crescente, perché il sistema consente in questo caso solo l'ordinamento alfabetico crescente. Nel post precedentemente analizzato, la stringa di classificazione rispetta l'ordine di citazione standard perché la categoria 'Entità' che deve precedere la categoria 'Agenti' ha come equivalente notazionale la lettera 'B' che viene prima della lettera 'M'.

In uno schema di classificazione basato sul principio delle faccette deve essere prevista anche una ‘**tavola invertita**’ in cui sia applicato il ‘**principio di inversione**’. Ciò vuol dire che nel momento in cui vado ad elencare tutti i post dell’archivio nei quali sono espresse faccette diverse conviene applicare il principio che *il generico preceda lo specifico* e quindi nella tavola invertita del blog <<http://www.biblioatipici.it/archives.shtml#cat>> le faccette sono elencate o tabulate in ordine rovesciato rispetto a quello di citazione.

Purtroppo il CMS consente l'ordinamento alfabetico inverso solo attraverso tag legati alla preparazione degli archivi e quindi solo nella tavola invertita, non nella stringa di classificazione dei singoli oggetti. Questo perché la tavola invertita e' un prodotto dell'archivio del blog e usa dei tag legati all'archivio (si veda il manuale del software on-line

<[http://www.sixapart.com/movabletype/docs/3.2/a\\_template\\_tag\\_reference/subcategory/](http://www.sixapart.com/movabletype/docs/3.2/a_template_tag_reference/subcategory/)>) mentre la stringa di classificazione, non essendo un prodotto d'archivio, è legata al post e quindi usa dei tag di tipo diverso <[http://www.sixapart.com/movabletype/docs/3.2/a\\_template\\_tag\\_reference/entry/](http://www.sixapart.com/movabletype/docs/3.2/a_template_tag_reference/entry/)>. Con i tag dell’archivio c’è la possibilità di modificare il sort, scegliendo l'ordinamento inverso (<MTSubCategories sort\_order="descend">), invece con i tag dei post non è possibile e si deve per forza adottare un ordinamento crescente <[http://www.sixapart.com/movabletype/docs/3.2/a\\_template\\_tag\\_reference/subcategory/#entry-6090](http://www.sixapart.com/movabletype/docs/3.2/a_template_tag_reference/subcategory/#entry-6090)>.

Proprio per questo motivo, la notazione è studiata affinché sia diretta all'interno delle stringhe; in questo modo saranno le categorie standard meno concrete (Tempo, Spazio) ad avere le lettere più alte (T: Tempo; S: Spazio), piuttosto che il contrario come sarebbe logico e come succede nella Colon classification o nella Classificazione Bliss.

Se si clicca sulla faccetta ‘isole’ nella tavola invertita, ad esempio, si apre un collegamento e vengono visualizzati tutti i post che nella stringa di classificazione hanno la faccetta ‘isole’. L’ordine di questi post è alfabetico per titoli mentre, a rigor di precisione, dovrebbe essere un ordine che segue il principio di inversione e cioè un ordinamento alfabetico decrescente delle faccette: quindi prima il post la cui stringa di

classificazione inizia con la faccetta 'L', poi quelli con la faccetta 'H' e infine quello con la faccetta 'B'. Questo è un limite imposto dal CMS utilizzato.

Un altro limite del sistema è che esso consente di ricercare i post per categoria soltanto con la funzione '*browse*' (navigazione delle notizie per argomento) e non invece anche con la funzione '*search*'. Questo comporta che, ad esempio, non sia possibile utilizzare due o più faccette come attributi di ricerca in combinazione.

Infine, un'altra possibilità offerta dall'organizzazione in categorie del blog, è quella di generare **RSS feed** specifici **per argomento**, oltre a quelli generali.

Un *feed* è l'elenco di un certo numero di news di un sito in linguaggio XML; un *feed RSS* è messo a disposizione da un fornitore di contenuti informativi ed è pronto per essere utilizzato da altri. Per utilizzare i feed, occorre un programma *aggregatore*, cioè un software in grado di seguire gli aggiornamenti di più *feed RSS* e di renderli visualizzabili all'utente. L'utente indica direttamente all'*aggregatore* i *feed RSS* che intende seguire.

In Biblioatipici i *feed RSS* sono aggiornati ad ogni nuovo post e possono contenere anche i commenti relativi ai post (ma in questo blog non li contengono). Tutti i *feed RSS* disponibili sono raggiungibili a partire dalla URL di raggruppamento <[http://www.biblioatipici.it/2005/01/feed\\_rss\\_di\\_bib.shtml](http://www.biblioatipici.it/2005/01/feed_rss_di_bib.shtml)>.

Per concludere si può dire che l'omogeneità semantica dei post contenuti in questo blog, rispetto al blog precedente, rende particolarmente azzeccata la scelta di questa classificazione rispetto per esempio alla CDD: nelle tavole di classificazione, infatti, non occorre prevedere tutte le possibili combinazioni di concetti, a differenza di quanto avviene nelle classificazioni tradizionali, che sono dette per questo *enumerative*. È sufficiente elencare gli isolati che possono comparire in ciascuna faccetta e le regole per combinarli.

I vantaggi dell'applicazione a faccette sono molteplici sia per chi gestisce il sito sia per l'utente. Per citarne alcuni, si può dire che con le faccette si assicura anche all'indicizzatore la possibilità di generare le classi di cui ha bisogno in fase di indicizzazione; con un numero poco elevato di faccette e di fuochi si ottiene un numero di classi esponenziale. Infine tra i vantaggi per l'utente, lo schema a faccette consente

accessi multipli sulla base dei diversi bisogni informativi, oltre che risultare coerente e intuitivo.

## 2. CLASSIFICAZIONE DECIMALE DEWEY

**Atipicheletture** <<http://www.biblioatipici.it/letture/>> è un *weblog* che nasce nel dicembre del 2005 ed è strettamente legato al blog Biblioatipici: in esso, infatti, sono segnalate le letture del gruppo Biblioatipici (ma vorrebbe aprirsi anche alla collaborazione di persone esterne a questo gruppo). Una specie di “*diario di lettura collettivo*”.

Ogni post corrisponde alla segnalazione di un libro che può riguardare il mondo del lavoro e i temi ‘atipici’, può essere di ambito LIS, ma può anche semplicemente essere un libro di narrativa italiana o straniera, di sociologia, di storia, insomma afferente a qualsiasi ambito disciplinare.

Ad oggi il blog contiene quasi **70** post e sta assumendo una fisionomia sempre più definita, con connotati tipicamente biblioteconomici: la descrizione dei libri è redatta secondo lo standard ISBD(M), ogni post è classificato con la CDD (Classificazione Decimale Dewey); inoltre per ogni post è attivo un link che localizza il documento nell’OPAC dell’Indice SBN <<http://opac.sbn.it/cgi-bin/IccuForm.pl?form=WebFrame>>.

Per quanto riguarda questi aspetti, al nascere del blog, ci fu nella *mailing list* Biblioatipici un simpatico dibattito su quanto fosse auspicabile impostare il sito secondo criteri biblioteconomici. Si formarono due correnti di pensiero: da una parte chi riteneva che bisognasse essere meno ‘bibliotecari’ e quindi personalizzare i post con spunti originali, commenti, brani, richiami da un libro all’altro in un contesto molto libero e contaminato; dall’altra chi era convinto che il *weblog* potesse diventare anche una palestra per strutturare biblioteconomicamente l’architettura del sito. Gli archivi di gennaio 2006 della *mailing list* in cui è possibile seguire il dibattito si trovano qui: <<http://lists.biblioatipici.it/mailman/private/ml/2006-January/thread.html>>.

A favore della prima tesi c’era anche un’altra argomentazione e cioè che se l’inserimento dei post fosse diventato troppo macchinoso questo avrebbe scoraggiato molti a contribuire.

La soluzione fu che “*l’una cosa non escludeva l’altra*” e cioè che [con le parole di Livia Castelli, membro del gruppo Biblioatipici]:

*“Cosa c’è di più contaminante di un catalogo che da modo di afferrare il libro di cui si parla, e magari altri affini, nella sua concretezza, almeno come localizzazione, anziché continuare a parlarne senza magari mai leggerlo? Perché poi dobbiamo vedere la "personalizzazione" come qualcosa che esclude categoricamente gli strumenti di lavoro e organizzazione della conoscenza che ogni giorno utilizziamo e che dovremmo augurarci vengano diffusi in quanto utili e ricchi di potenzialità [...]? E più in generale: perché le nostre conoscenze, intuizioni, idee professionali dovrebbero essere isolate in un compartimento stagno quando diventiamo lettori, blogger o qualsiasi altra cosa? “*

Queste sono le caratteristiche di ciascun post:

- ogni segnalazione libraria è corredata dalla copertina del libro e dalla scheda ISBD relativa;
- segue quella che può essere o la trama del libro o una citazione, o una ‘suggerione’;
- grazie all’utilizzo di parole-chiave è possibile indicizzare i seguenti dati bibliografici: autore, editore, collana. In questo modo si possono immediatamente visualizzare tutti i titoli presenti nel blog dello stesso autore, dello stesso editore e con la stessa collana;
- è possibile localizzare il documento in SBN: cliccando infatti sul link ‘Localizza titolo in SBN’, predisposto per ogni post, si attiva una query su Indice SBN che imposta automaticamente il titolo del post (= titolo libro) come chiave di ricerca del titolo. Il risultato in alcuni casi è affetto da rumore, essendo una funzione di *search* solo per titolo e non incrociabile. Si veda ad esempio il seguente post: <<http://www.biblioatipici.it/letture/2006/01/piattaforma.shtml>>, in cui il titolo del libro risulta essere poco significativo per questa ricerca;
- L’indicazione dell’ISBN tra le keywords, nella seguente forma ‘ISBN=XXXXXXXXXXXX’ (senza spazi tra i gruppi di cifre) permette l’attivazione automatica della ricerca del libro per l’acquisto in alcune librerie on-line selezionate (BOL, IBS, DEAstore);

- Per ogni post è possibile lasciare un commento;
- Tra i dati indicizzati vi è anche il 'lettore' (che poi è colui che inserisce il post) e quindi è possibile vedere tutti i libri letti da un lettore

### Classificazione Decimale Dewey

La scelta di utilizzare la CDD per classificare i post nasce dalla connotazione stessa del blog: *Atipicheletture* è un blog di segnalazioni librarie relative agli ambiti disciplinari più vari, pertanto ogni post corrisponde a un libro, proprio come se ci si trovasse di fronte allo scaffale di una biblioteca multidisciplinare. In un tale contesto la CDD costituisce uno standard di classificazione consolidato e l'adozione di uno schema diverso rischierebbe di disorientare, anche perché questo è un blog gestito da bibliotecari per altri bibliotecari.

Assegnando a ciascun post (=libro) una stringa di classificazione, sarà possibile recuperare tutti i libri dello stesso argomento nella stessa classe e quelli di argomento simile nelle classi attigue (proprio come in uno scaffale aperto). Si veda, ad esempio l'archivio della classe 853.914 (Narrativa italiana, 1945-): <http://www.biblioatipici.it/letture/800/850/853914/>.

Inoltre le funzionalità che la CDD offre se applicata a un CMS superano la fisicità di uno scaffale. È infatti possibile navigare da una classe nella classe immediatamente superiore, cioè meno specifica: se ad esempio un post è stato classificato come narrativa italiana del 21. secolo, con un semplice link è possibile visualizzare tutti i post classificati nella narrativa italiana di tutti i tempi (senza specificazioni temporali).

La navigazione è facilitata anche dalla presenza di un catalogo classificato <http://www.biblioatipici.it/letture/cdd.shtml>, nel quale sono visibili le classi Dewey utilizzate per classificare i post, rappresentate da una *notazione* e una *vedetta* (= equivalente verbale della notazione), e il link ai post inseriti. La notazione della CDD ha caratteristiche abbastanza mnemoniche: è sulla base di 10 numeri; l'intero scibile è inteso come l'unità, ripartita in 10 classi principali, ciascuna delle quali ha 10 divisioni ed ognuna di queste 10 sezioni e così via per espansioni successive indicate esclusivamente con numeri. I numeri si leggono come decimali, come se fossero preceduti da 0 più

virgola: non dunque come numeri interi, ma cifra per cifra (quindi 813.54 precederà 813.6). Si tratta chiaramente di una notazione *espressiva* perché a simboli più lunghi corrispondono classi più specifiche.

Si può dire che questo catalogo classificato, così come è concepito, con l'ausilio delle *vedette*, costituisca una sorta di *Dewey browser* nel senso che permette di utilizzare la CDD come strumento di *retrieval* da parte degli utenti (anche se poco pratici di CDD).

Dalla tavola risulta evidente che la CDD, essendo una classificazione *gerarchico-enumerativa*, prevede tutte le possibili combinazioni di concetti e quindi, rispetto alla classificazione a faccette che è *analitico-sintetica*, si espande maggiormente perché giunge ad indicare soggetti molto specifici. Per far questo la CDD utilizza numerosi livelli gerarchici, sfruttando al massimo la caratteristica delle categorie del CMS di essere infinitamente suddivisibili. Al contrario, la classificazione a faccette utilizza in genere pochi livelli gerarchici, infatti nel blog Biblioatipici ne troveremo 3 in tutto.

In questo blog, invece, anche se i post sono ancora poco numerosi e quindi la CDD si è poco espansa al momento, siamo già a **4 livelli gerarchici**. Si veda ad esempio la classe 000 in cui abbiamo 020 figlio di 000, un 025 figlio di 020, 025.3 figlio di 025 e così via...

Il catalogo classificato del blog è opportunamente ospitato dal CMS all'interno di una *struttura ad albero*, nella quale ci sono gerarchie e percorsi obbligati: la navigazione assume la forma di un percorso a gradini, dalla categoria padre alla categoria figlio e così via.

Nonostante la CDD si adatti piuttosto efficacemente a questo blog, come abbiamo visto, tuttavia, in generale, le applicazioni al Web che ne sono state fatte risultano alquanto insoddisfacenti, soprattutto se paragonate alle funzionalità di uno schema basato sul principio delle faccette.

Alcuni svantaggi dell'adozione della CDD, evidenti anche in questo sito, sono:

1. la presenza di un elevato numero di livelli gerarchici impedisce l'immediatezza della navigazione: il numero delle opzioni di un menu, se troppo elevato, influenza, infatti, l'usabilità dello stesso e produce sovraffollamento informativo;
2. è uno schema poco scalabile: si adatta con difficoltà a eventuali aggiunte di componenti o a parziali modifiche nella sua struttura;
3. è strutturalmente chiuso e conservativo: non consente l'inserimento da parte del classificatore di una nuova categoria; solo il compilatore può modificare la classificazione, redigendo un foglio di revisione ufficiale o pubblicando un'edizione successiva dello schema.

### 3. JITA SCHEMA

**Il blog di Andrea Marchitelli** <<http://www.biblioatipici.it/andrea/>> è un weblog personale. Contiene post che rientrano nelle discipline LIS (Library and Information Science): segnalazioni di incontri professionali, novità nel panorama bibliotecario italiano e straniero, pubblicazioni di interesse scientifico nonché tutte le informazioni legate all'attività professionale dell'autore del blog.

Ad oggi il blog non contiene ancora un elevato numero di post ed ha un bacino di utenti non particolarmente ampio se confrontato con i blog successivi.

Le tecniche di *retrieval* abilitate in questo blog sono: la funzionalità 'search' che cerca nel *full text* dei post, lo scorrimento per mese di pubblicazione, lo scorrimento per categorie e, infine, l'impiego dei tag.

Considerato l'ambito disciplinare di questo blog, lo schema di classificazione adottato è lo **schema JITA** (acronimo dei nomi degli autori dello schema). È uno schema di ambito LIS: nasce infatti per classificare i documenti di E-LIS (E-prints in Library and Information Science) <<http://eprints.rclis.org/>>, l'archivio internazionale ad accesso aperto per la biblioteconomia. È utilizzato anche dalla Biblioteca dell'Area di Ricerca CNR di Bologna <<http://biblio-eprints.bo.cnr.it/view/subjects/>>.

Nella *tavola* <<http://www.biblioatipici.it/andrea/jita.shtml>> i soggetti sono elencati in maniera sistematica, così da rendere visibili le relazioni tra di loro. È uno schema classificatorio molto semplice, fusione e rielaborazione della NewsAgentTopic Classification Scheme (mantenuta da Mike Keen at Aberystwyth, UK, fino al 31 marzo 1998) e della RIS classification scheme (ora cessata) Review of Information Science originariamente concepita da Donald Soergel (University of Maryland). È gerarchico su due livelli (il secondo livello è stato aperto solo nel 2005) e ad oggi contiene oltre 120 *item*. I documenti possono essere classificati con uno o più *item*.

Lo JITA schema è diviso in 12 blocchi (A-L) creati attorno a 3 aree tematiche, non manifeste:

1. **teorica e generale:** include aspetti teorici e generali della biblioteconomia e dell'informazione, uso dell'informazione e sociologia dell'informazione
2. **funzioni dell'utente, direzionali e gestionali: livello intermedio (incluse questioni socio-economiche e legali):** si trovano qui lavori diretti agli utenti, l'alfabetizzazione e la lettura; biblioteche e repository di informazioni; temi dell'editoria e legali, compreso il copyright e la gestione dei diritti, gestione per progetti e industria, professione e educazione.
3. **oggetti, temi pragmatici e questioni tecniche a livello specifiche:** riguarda fonti, supporti e canali d'informazione; il trattamento delle informazioni per i servizi informativi, i servizi tecnici in biblioteca, archivi e musei, tecnologie dell'informazione e 'tecnologia bibliotecaria'.

La *notazione* è composta da lettere dell'alfabeto latino in carattere maiuscolo e rende immediatamente visibile, anche ad un utente non esperto, l'ordine sotteso alla struttura dello schema classificatorio.

Lo scopo di questo schema è essenzialmente pragmatico e cioè presentare in visualizzazione e recupero, attraverso modalità di *browsing*, i documenti suddivisi in ampie categorie senza scendere nel dettaglio (i livelli gerarchici sono solo due).

Anche se si tratta di uno schema molto semplice, esso rientra tra gli schemi di classificazione *gerarchico-enumerativa* o *monodimensionali* perché la relazione

tra i soggetti nella tavola è di tipo gerarchico e perché lo schema prevede l'enumerazione di tutti gli eventuali soggetti semplici e composti.

Gli svantaggi di questo schema sono molteplici: in primo luogo è stato pensato per classificare i documenti presenti in E-LIS, dunque è poco adattabile al di fuori di questo ambito; in secondo luogo è uno schema conservativo perché strutturalmente chiuso, non consente, infatti, l'inserimento da parte del classificatore di nuove categorie; infine, non sono previsti aggiornamenti periodici da parte del comitato editoriale, come accade invece per la CDD.

Nonostante ciò, data la sua semplicità e il settore disciplinare rappresentato, in questo blog può trovare un efficace utilizzo, considerato anche che si tratta di uno schema sufficientemente *coerente*.

Nel blog in esame l'archivio di ogni categoria di post è collegato dinamicamente con il suo omologo in E-LIS (si può osservare qui: <http://www.biblioatipici.it/andrea/i/id/index.shtml>). Inoltre nella home page, sono elencati e linkabili gli ultimi paper pubblicati in E-LIS, grazie all'uso dei feed RSS, resi disponibili nell'archivio E-LIS <<http://eprints.relis.org/last.xml>>.

Infine, nel blog è attiva un'altra funzionalità di *indicizzazione semantica* che consente di individuare per ciascun post del blog uno o più **tag** <<http://www.biblioatipici.it/andrea/tags.shtml>>. I tag costituiscono, in questa accezione, il soggetto/argomento del post, espresso con parole chiave, dunque in senso più generico e meno gerarchico rispetto alla classe. Mediante dei servizi di '*social tagging*', come **Technorati** <<http://www.technorati.com/>>, un motore di ricerca, attraverso i tag, aggrega i post di vari blog: cliccando quindi su un certo tag del blog si scatena una ricerca in Technorati e si trovano i blog che citano lo stesso tag e dunque di argomento simile. Tutti i tag individuati si trovano sullo stesso livello: non esiste una gerarchia (tag-padri e tag-figli); l'unico fattore distintivo è la loro popolarità: alcuni tag sono utilizzati molto spesso per descrivere un item, altri molto raramente.

## Bibliografia

- Broughton, V. 2004. *Essential classification*. London: Facet Publications.
- De Robbio, A., Subirats-Coll, I. 2005. E-LIS : an International Open Archive Towards Building Open Digital Libraries. *HEP Libraries Webzine*, 11.  
<<http://library.cern.ch/HEPLW/11/papers/1/>>
- Foskett, A.C. 1996. *The subject approach to information*, 5th ed. London: Library association.
- Gnoli, C. 2004. *Classificazione a faccette*. Roma: AIB.
- Gnoli, C., Marino, V., Rosati, L. 2006. *Organizzare la conoscenza*. Milano: Tecniche nuove.
- Hammond, T., Hannay, T., Lund, B., Scott, J. 2005. Social Bookmarking tools (I): a general review. *D-Lib Magazine*, 11 (4).  
<<http://www.dlib.org/dlib/april05/hammond/04hammond.html>>
- Kepner, L.T. 2002. Workable solutions with Dewey classification. *Technicalities*; 22 (4). 3, 9, 12-13.
- Mai, J.E. 2004. Classification of the web: challenges and inquiries. *Knowledge Organization*, 31 (2). 92-97.
- Ranganathan, S.R., 1967. *Prolegomena to library classification*. 3. ed. - Bangalore : Sarada Ranganathan endowment for library science, 1967.
- Slavic, A. 2002. *Facet analytical theory in managing knowledge structure for humanities*. University College London <<http://www.ucl.ac.uk/fatks/>>.
- Vickery, B.C. 1960. *Faceted classification: a guide to the construction and use of special schemes*. London: Aslib.
- Vizine-Goetz, D., Hickey, T. 2006. Getting visual with DeweyBrowser. *NextSpace*, n. 1. 18-19.

	<b>Biblioatipici schema</b>	<b>DDC</b>	<b>JITA</b>
<b>Nome del blog</b>	Biblioatipici	AtipicheLecture	Andrea Marchitelli's blog
<b>Numero di post</b>	495	85	112
<b>Home made</b>	SI A faccette	NO Gerarchico	NO Gerarchico
<b>Numero di classi / FOCI</b>	61	67	135 <sup>(4)</sup>
<b>URL dello schema</b>	<a href="http://www.biblioatipici.it/archives.shtml#cat">http://www.biblioatipici.it/archives.shtml#cat</a>	<a href="http://www.biblioatipici.it/letture/cdd.shtml">http://www.biblioatipici.it/letture/cdd.shtml</a>	<a href="http://www.biblioatipici.it/andrea/jita.shtml">http://www.biblioatipici.it/andrea/jita.shtml</a>
<b>Esempio di stringa di classificazione</b>	<i>H11: Concorsi non scaduti::J94: Tempo indeterminato::M58: Università::S51: Nord Italia</i>	<i>346.0482 (Diritto della proprietà. Diritto d'autore - Copyright)</i>	<i>LJ (Software)</i>
<b>Utilizzati anche da</b>	- - -	BUBL Information service < <a href="http://bubl.ac.uk/">http://bubl.ac.uk/</a> >  Canadian information by subject < <a href="http://www.collectionscanada.ca/caninfo/esub.htm">http://www.collectionscanada.ca/caninfo/esub.htm</a> >	E-LIS < <a href="http://eprints.rclis.org/view/subjects/">http://eprints.rclis.org/view/subjects/</a> >

<sup>4</sup> Questo è il numero complessivo delle classi dello schema. Nel blog ne sono state utilizzate meno di 50, per ora.

