

## Dalla classificazione per i cittadini alla classificazione *dei* cittadini Il caso *TaggaTO* \*

FRANCO CARCILLO - LUCA ROSATI

*Il progetto TaggaTO, un progetto web sviluppato dal Comune di Torino, intende migliorare l'esperienza degli utenti all'interno del sito del Comune <<http://www.comune.torino.it>> integrando una tassonomia standard sviluppata centralmente (top-down) con una classificazione mediante tag creata direttamente dagli utenti dal basso (bottom-up). Lo schema di classificazione top-down è stato progettato seguendo lo standard anglosassone IPSV (Integrated Public Sector Vocabulary, schema di classificazione conforme all'ISO 2788), raffinato empiricamente mediante test con utenti e analisi dei file di log. Lo schema bottom-up funziona come un sistema di tagging sociale (folksonomy). Quest'ultimo non è semplicemente giustapposto al primo, ma integrato con esso in modo da ottenere un sistema coerente.*

Parole chiave: eGovernment - Metadati - Architettura dell'informazione - Classificazione - Tassonomie - Folksonomie - IPSV [Integrated Public Sector Vocabulary] - Torino - Taggato

### Introduzione: il meglio dei due mondi

La classificazione sociale mediante *tagging* collaborativo (o *folksonomy* <sup>1</sup>) può essere vista non solo come un'alternativa ai modelli di classificazione tradizionali messi a punto in ambito bibliotecario, ma anche come uno strumento complementare da utilizzare accanto a questi ultimi. In questo modo, è possibile correggere i limiti che ciascuno schema di classificazione inevitabilmente possiede e rafforzarne nel contempo i pregi. Considerando l'ampio spettro di esigenze e modelli mentali che un sito di *e-government* deve soddisfare, l'impiego di un modello di classificazione "misto" può portare notevoli benefici.

\* Sebbene l'articolo sia frutto di un lavoro comune, Franco Carcillo ha scritto i paragrafi *Introduzione, Tassonomie e folksonomy: pro e contro*; Luca Rosati gli altri. L'articolo è la traduzione, con poche variazioni, di Carcillo - Rosati, 2007.

<sup>1</sup> Per maggiori dettagli sul fenomeno del *tagging* collaborativo o *folksonomy*, vedi par. *Tassonomie e folksonomy: pro e contro*.

Così, l'obiettivo del progetto qui descritto (realizzato nel corso del 2007 dal Comune di Torino per il sito web <<http://www.comune.torino.it>>) è mostrare come lo spazio piatto di parole chiave (*tag*) creato dagli utenti dal basso possa essere efficacemente fuso con un sistema di classificazione poli-gerarchico, così da migliorare l'architettura dell'informazione del sito. Oltre a creare un ordine e una gerarchia all'interno della massa piatta dei *tag*, l'approccio misto tassonomia-*folksonomy* è in grado di rafforzare due aspetti chiave del sistema: il profumo dell'informazione (*information scent* - Pirolli, 2007) e la raccolta delle bacche (*berypicking* - Bates, 1989)<sup>2</sup>. Ogni architettura informativa possiede infatti due assi fondamentali:

- un asse verticale (o paradigmatico), che individua le relazioni gerarchiche che ogni elemento del sistema intrattiene con gli altri;
- un asse orizzontale (o sintagmatico), che individua le relazioni di contiguità semantica fra gli elementi del sistema (Figura 1).

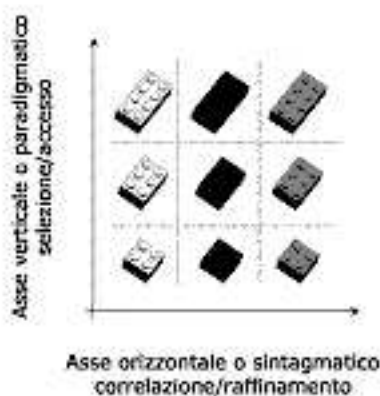


Figura 1 - I due assi dell'architettura dell'informazione

<sup>2</sup> Il termine "raccolta delle bacche" (*berypicking*) applicato al settore dell'*information retrieval* è stato coniato da Marcia Bates (1989) e indica una fondamentale modalità con cui l'uomo cerca le informazioni (in ogni ambiente, sia fisico sia digitale) modificando continuamente la propria strategia a seconda dei nuovi ritrovamenti; per questo motivo la Bates parla anche di "ricerca evolutiva". Il concetto di "profumo dell'informazione" (*information scent*), è stato coniato da Pirolli - Card, 1999 nell'ambito della teoria dell'*information foraging*, sviluppata applicando concetti propri della biologia (strategie comportamentali degli animali per la ricerca del cibo) al campo dell'*information retrieval*. Il profumo dell'informazione indica un insieme di indizi che suggeriscono il probabile valore di rilevanza delle informazioni a cui l'utente si sta avvicinando nel corso della sua ricerca. Questo concetto si lega a quello di raccolta delle bacche, dato che il "profumo" trasmesso da una risorsa influenza l'evoluzione del percorso dell'utente: l'utente si dirige verso le informazioni con più alto

Nel nostro caso, la combinazione di *folksonomy* e tassonomia permette una migliore gestione di entrambi questi assi:

dal punto di vista verticale o paradigmatico, quando un utente (per salvare una pagina nei propri preferiti) associa un *tag* a una categoria della tassonomia, il sistema suggerisce *tag* simili appartenenti alla medesima categoria o usati in associazione con quello corrente (così come fa Delicious <<http://del.icio.us/>>);

dal punto di vista orizzontale o sintagmatico, il sistema permette di visualizzare tutti gli altri *tag* appartenenti alla stessa categoria, o usati dallo stesso utente.

Altre caratteristiche principali sono:

sistema di suggerimento dei *tag* durante l'inserimento dei *tag*, o durante la digitazione nel motore di ricerca interno  
ricerca ibrida mediante *tag* e parole chiave presenti nella pagina  
possibilità di usare *tag* composti di più parole  
*tag history*.

### Tassonomie e *folksonomy*: pro e contro

«La mia ipotesi è che noi abbiamo una teoria popolare della classificazione stessa. Essa sostiene che le cose siano di per sé organizzate in tipi ben definiti, che i tipi siano caratterizzati da proprietà condivise, e che ci sia una sola tassonomia corretta dei tipi. È più facile mostrare cosa ci sia di errato in una teoria scientifica della classificazione che non in una teoria popolare. Una teoria popolare definisce il senso comune stesso. Quando una teoria popolare e una teoria tecnico-scientifica convergono, diventa ancora più difficile stabilire dove la teoria fa acqua - o se essa si possa effettivamente definire una teoria». (Lakoff, 1987, p. 121, citato in Bowker - Star, 1999, p. 33; traduzione nostra)

Le tassonomie sono sistemi di classificazioni *top-down*, creati da personale specializzato, quindi piuttosto centralizzati e precisi: si avvicinano cioè proprio a quel genere di classificazione che Lakoff definisce "teoria scientifica", e che riflette una logica di derivazione aristotelica<sup>3</sup>.

tasso di *scent*, tanto che una diminuzione significativa di "profumo" può determinare l'abbandono del percorso.

<sup>3</sup> Una classificazione aristotelica funziona sulla base di proprietà binarie che l'oggetto da classificare presenta o non presenta. Così, ad ogni livello della classificazione, si sceglie un insieme di queste proprietà per determinare l'appartenenza o non appartenenza dell'oggetto a una e una sola classe; un poligono appartiene pertanto alla classe dei rettangoli se: presenta quattro lati; i lati sono ortogonali fra loro; i lati opposti sono uguali.

All'opposto, le *folksonomy* sono sistemi di classificazione *bottom-up*, creati dagli utenti stessi mediante l'aggiunta di parole-chiave; sono sistemi distribuiti e quindi poco precisi, ma capaci di riflettere meglio il punto di vista degli utenti: si avvicinano perciò al modello di classificazione che Lakoff definisce "popolare", un modello assimilabile in questo caso alla teoria dei prototipi di Eleanor Rosch (più che ad Aristotele).

In termini filosofici, tuttavia, nessun sistema di classificazione può definirsi totalmente scientifico o totalmente empirico (popolare): qualunque classificazione, anche quando nasca come scientifica, tende sempre a modificarsi in senso popolare, in base cioè alle esigenze empiriche che l'uso inevitabilmente impone.

«Perché in alcune circostanze le nostre classificazioni appaiono nella pratica come prototipiche anche se in principio esse erano aristoteliche? Per due principali ragioni: perché ogni sistema di classificazione è legato a un insieme di pratiche di codifica e perché i sistemi di classificazione in genere (non stiamo dando un giudizio *ex cathedra*) riflettono inevitabilmente le motivazioni conflittuali e contraddittorie del contesto socio-etnico da cui derivano [...].

In aggiunta a questa eredità, vi è un vero e proprio rasoio di Occam. Quando i medici devono indicare le cause di un decesso essi si trovano a confrontarsi con una serie di difficili valutazioni (che richiederebbero un'autopsia o ulteriori indagini diagnostiche). I medici scelgono semplicemente la via più breve, classificando spesso le cause del decesso sotto la generica categoria "altro". Essi possono quindi tornare a occuparsi dei loro pazienti vivi [...]. Così, la bellezza della classificazione aristotelica lascia spazio a un sistema di classificazione molto più sfumato, che condivide tratti con classificazioni basate sul prototipo del buonsenso, vale a dire un insieme di oggetti eterogenei tenuti assieme da meccanismi analogici o metaforici [...].

Non vi è quindi una demarcazione netta fra classificazioni popolari e classificazioni scientifiche. Di seguito, esamineremo una particolare linea di frattura fra le due: linea che viene costantemente ridefinita e che muta la propria natura man mano che la placca dell'esperienza vissuta si incanala sotto la crosta della conoscenza scientifica». (Bowker - Star, 1999, p. 64, 65, 67)

Tutte queste considerazioni ci hanno condotto all'idea di combinare classificazione *top-down* e classificazione *bottom-up*, sia per ottenere (per così dire) il meglio dei due mondi, sia - soprattutto - per cercare di comprendere meglio, tramite la classificazione popolare, i modelli mentali dei cittadini e correggere di conseguenza i limiti della tassonomia *top-down*.

Piuttosto che in termini di opposizione, tassonomie e *folksonomy* possono allora essere viste come strumenti alternativi e complementari, che in virtù delle loro peculiarità si prestano ad essere utilizzati ciascuno in contesti e con obiettivi specifici.

### Anatomia delle tassonomie

Le tassonomie sono sistemi di organizzazione della conoscenza costituiti da termini disposti in un albero gerarchico. Esse sfruttano quindi sistemi di classificazione gerarchico-enumerativi da cui ereditano inevitabilmente pregi e difetti.

I sistemi ad albero permettono una notevole precisione e ordine nella classificazione degli elementi di una collezione, favorendo così un approccio mirato all'informazione (*known-item seeking*). Queste stesse caratteristiche costituiscono però anche i limiti di questi sistemi.

Essendo schemi di classificazione *top-down* (cioè creati e mantenuti da pochi esperti o da un'istituzione), le tassonomie sono sistemi piuttosto centralizzati, in cui eventuali modifiche possono essere introdotte solo dall'alto.

Il rigore dei modelli ad albero si traduce spesso in una eccessiva rigidità e chiusura del sistema: l'introduzione di nuove categorie in un punto del sistema spesso richiede una modifica anche di altre porzioni dello stesso, e può essere effettuata solo dal compilatore dello schema.

Essendo molto ramificati in profondità, i sistemi ad albero tendono spesso a nascondere le classi più popolari (quelle più usate dagli utenti) "spingendole" verso il basso.

La struttura gerarchica si adatta bene a strategie di tipo *known-item seeking* (le strategie messe in atto da utenti che sanno già cosa cercano), mentre non si adatta affatto a strategie di *exploratory-seeking* (utenti che non sanno bene cosa cercare, curiosi, utenti inesperti). Inoltre, essa non favorisce quel modello di navigazione definito da Bates (1989) "raccolta delle bacche" (vedi paragrafo iniziale), così tipico del web.

#### Anatomia delle *folksonomy*

La forza delle *folksonomy* non è la precisione, ma l'ampia adesione popolare: la gente, appunto. Anziché sforzarsi di creare sistemi di classificazione che incontrino i modelli mentali degli utenti, si fa in modo che siano gli utenti stessi - dal basso - a far emergere spontaneamente (attraverso un processo collaborativo e una sorta di selezione naturale) modelli mentali comuni, condivisi.

Sarebbe infatti più giusto definire le *folksonomy* come sistemi di classificazione distribuita. E, come tali, le *folksonomy* possono essere utilizzate anche come strumenti di analisi: per studiare i modelli mentali di ampie comunità, attraverso il modo in cui i loro appartenenti descrivono (classificano) una collezione di item; possono essere utilizzate come supporto per elaborare categorizzazioni, scegliere termini preferiti o allestire vocabolari controllati (Quintarelli, 2005).

I detrattori delle *folksonomy* citano questi difetti principali.

Le *folksonomy* mancano di precisione nella categorizzazione e nel linguaggio, con il rischio di proliferazione di molte varianti linguistiche per uno stesso concetto, e conseguente perdita di controllo sul vocabolario e sui sinonimi;

è pur vero che gli ultimi sistemi basati su *tag* prevedono oggi un sistema di suggerimenti capaci di mostrare i *tag* correlati a quello selezionato. Inoltre, come critica a questa posizione, possiamo ricordare la stessa citazione di Lakoff riportata sopra: un sistema di classificazione è sempre in qualche misura ibrido, poiché il suo fine ultimo è pratico (l'uso che se ne fa), esso inevitabilmente mescola una base logico-scientifica a criteri prettamente popolari, empirici.

I *tag* usati nelle *folksonomy* sono privi di una qualunque strutturazione gerarchica: manca del tutto una categorizzazione in categorie e sottocategorie, sia in senso gerarchico-enumerativo sia in senso analitico-sintetico. «Tags are a flat space of keywords». Questa obiezione - se pur vera nella sua essenza - non coglie l'aspetto più innovativo del fenomeno, che non consiste tanto in questa massa piatta di *tag* quanto nella loro aggregazione da parte degli utenti.

Le *folksonomy* possiedono un basso quoziente di trovabilità (*findability*); esse sono quindi più adatte a una ricerca di tipo esplorativo (*exploratory seeking*) piuttosto che a una ricerca mirata (*known-item seeking*).

Per quanto riguarda gli aspetti positivi o di interesse delle *folksonomy*, possiamo annoverare questi vantaggi.

Anzitutto, non tutti i limiti delle *folksonomy* sono difetti. Se da un lato c'è indubbiamente una perdita di precisione, questa perdita viene bilanciata dall'altro da una serie non irrilevante di benefici, in termini di suggerimenti sulla rappresentazione e l'uso da parte di grandi quantità di utenti di certi concetti, e quindi sui modelli mentali dei navigatori. In questo senso, le *folksonomy* possono essere viste non solo come sistemi di classificazione ma anche come strumenti di ausilio alla progettazione della classificazione stessa. Le *folksonomy* sono più idonee alla scoperta dell'informazione (*discoverability*) che alla ricerca mirata (*findability*). Rifacendoci ai modelli visti, potremmo dire che le *folksonomy* favoriscono un modello di ricerca del tipo "raccolta delle bacche", integrando fortemente al loro interno il concetto di serendipità<sup>4</sup>.

<sup>4</sup> Serendipità è un neologismo ancora poco usato nella lingua italiana mentre è assai più diffuso nel mondo anglosassone. Proviene infatti dal vocabolo inglese *serendipity*, parola coniata nel 1754 dal letterato Horace Walpole il quale, rimanendo colpito dal racconto de *I tre principi di Serendippo* di Cristoforo Armeno, ne estrasse un personalissimo principio. Serendipità è dunque, filosoficamente, lo scoprire una cosa non cercata e imprevista mentre se ne sta cercando un'altra. Ma il termine non indica solo fortuna: per cogliere l'indizio che porterà alla scoperta occorre essere aperti alla ricerca e attenti a riconoscere il valore di esperienze che non corrispondono alle originarie aspettative (Fonte: *Wikipedia*).

Economicità: i sistemi basati su *tag* sono estremamente economici da implementare (sia in termini di tempo sia in termini di denaro), dato che tutto il lavoro di classificazione è fatto dagli utenti *a posteriori*.

## Il progetto *TaggaTO*: componenti del sistema di classificazione

La componente tassonomica: IPSV [Integrated Public Sector Vocabulary]

Per raffinare la tassonomia del sito del Comune di Torino abbiamo operato un confronto con lo standard britannico IPVS - Integrated Public Sector Vocabulary (UK Cabinet Office eGovernment Unit, 2006). Lo IPSV è un vocabolario controllato (definito anche come schema di codifica - *encoding scheme*) utilizzato come standard di classificazione per tutta la documentazione elettronica della pubblica amministrazione britannica, in modo da facilitarne il reperimento sia dai cittadini sia dagli addetti. Questo modello è conforme agli standard ISO 2788 e BS 8723, gli standard internazionale e britannico rispettivamente per i thesauri monolingui.

Lo IPSV è stato sviluppato per iniziativa del Department for Communities and Local Government (DCLG) e della e-Government Unit (e-GU) del Cabinet Office del Governo britannico. La qualità del progetto dipende anche da un approccio olistico al problema della classificazione nella pubblica amministrazione: al modello ha infatti lavorato un *pool* di esperti afferenti a diversi campi (linguistica, biblioteconomia, informatica).

Per il *redesign* della tassonomia del sito del Comune di Torino abbiamo pianificato tre fasi:

inventario dei contenuti: mappatura dei primi due livelli di navigazione del sito del Comune di Torino <<http://www.comune.torino.it>> (Figura 2);

comparazione fra la tassonomia attuale (quella emersa dall'inventario), lo standard IPSV e l'applicazione di tale standard al sito Directgov <<http://www.direct.gov.uk>> (Figure 3 e 4) per raffinare/correggere la tassonomia del sito del Comune di Torino;

una successiva e più profonda comparazione fra le etichette dei tre sistemi.



Figura 2 - La homepage del sito del Comune di Torino



Figura 3 - Lo standard IPSV - Integrated Public Sector Vocabulary, un thesaurus progettato e utilizzato dalla pubblica amministrazione anglosassone (UK Cabinet Office eGovernment Unit, 2006)





Figura 4 - Il sito britannico dei servizi al cittadino Directgov utilizza lo standard IPSV - Integrated Public Sector Vocabulary

Una sintesi della comparazione è mostrata nella Tabella 1 qui sotto (non abbiamo abbastanza spazio per illustrare nel dettaglio il processo di *redesign*).

Tabella 1 - Comparazione fra la tassonomia attuale del sito del Comune di Torino, lo standard IPSV e la tassonomia del sito britannico per i cittadini Direct.gov <<http://www.direct.gov.uk/>>.

Turin website	IPSV	Direct.gov.uk
Ambiente e verde	Environment	Environment and greener living
Casa e territorio	Housing	Home and community
Commercio e impresa	Business and industry	Information for businesses [in the footer menu]
Cultura, sport e tempo libero	Leisure and culture	Young people > Leisure
Formazione e Scuola	Education and skills	Education and learning
Lavoro e Orientamento	Employment, jobs and careers	Employment
Sanità e Servizi Sociali	Health, well-being and care	Health and well-being
Sicurezza ed emergenze	Public order, justice and rights > Security   Emergencies	In: Newsroom > Public safety [utilities menu]
Soldi e risparmio	In: Economics and finance > Personal finance	Money, tax and benefits
Tasse e agevolazioni fiscali	In: Economics and finance > Tax	Money, tax and benefits
Trasporti e viabilità	Transport and infrastructure	Travel and transport
Elemento Jolly (es. Speciale Giochi olimpici 2006)	--	--

La componente di *folksonomy*

Il modulo di *folksonomy* funziona come uno strato veloce di classificazione e navigazione sovrapposto allo strato più profondo costituito dalla tassonomia.

Due sono gli obiettivi principali di questo sistema:

- permettere agli utenti di salvare in un'area riservata (*Preferiti* o *Bookmarks*) pagine di interesse del sito del Comune di Torino, etichettandole mediante *tag*, così da non doverle ricercare ogni volta; i *tag* fungono in questo senso da scorciatoie di accesso per utenti abituali;
- permettere anche agli utenti che non hanno salvato *Preferiti* di usufruire dei *tag* creati dagli altri utenti come strumento di navigazione complementare alla tassonomia e al motore di ricerca.

Già dalla *home*, l'utente può in tal modo usare i *tag* sia come strumenti di accesso primario ai contenuti sia come strumento per raffinare ricerche effettuate inizialmente mediante la tassonomia standard o il motore di ricerca. I *tag* creati dagli utenti verranno salvati in un'area personale del sito, che ciascun utente potrà decidere in che misura rendere pubblica: del tutto, in parte o per nulla.

Le due modalità d'uso corrispondono ad altrettanti livelli di interattività:

- un uso più attivo, il primo
- un uso più passivo, il secondo.

Il primo è concepito per utenti più inclini alla partecipazione attiva o già abituati a usare sistemi simili (ad es. Del.icio.us), il secondo per coloro che, pur avendo un atteggiamento più conservativo, possono essere stimolati a modificare il proprio atteggiamento usufruendo comunque del lavoro svolto da altri utenti.

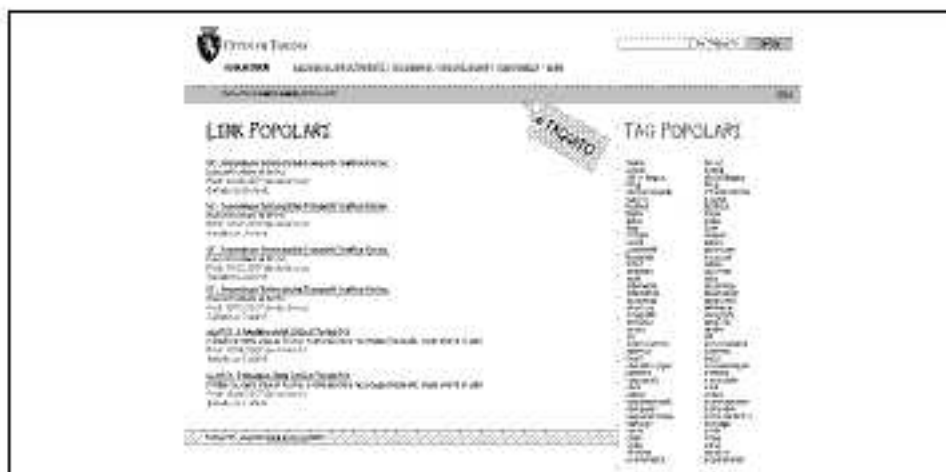


Figura 5 - La homepage di TaggaTO <<http://www.comune.torino.it/taggato/>>

## Unificazione

Abbiamo detto che il modulo di *folksonomy* non è semplicemente giustapposto al modulo della tassonomia, ma integrato con quest'ultimo. La sintesi avviene nella fase di *tagging*: durante il processo di inserimento il sistema chiede agli utenti di associare i *tag* alle categorie della tassonomia del sito, in modo da creare un raccordo fra i due strati di classificazione. Uno stesso *tag* potrà essere associato a più categorie contemporaneamente.

L'integrazione comporta dei vantaggi ulteriori in termini di funzionalità del sistema:

- disponibilità per l'utente di due sistemi di accesso complementari: mediante tassonomia e mediante *tag*;
- possibilità di interscambio fra i due sistemi: passaggio dalla tassonomia ai *tag* e viceversa, senza soluzione di continuità.

Più in dettaglio, ciò consente di:

- mostrare (tutti) gli altri *tag* appartenenti a una data categoria, quando l'utente associa un *tag* a quella categoria;
- ovviare al problema della piattezza dei *tag* (etichette su un unico indistinto livello), consentendo raggruppamenti per macro-categorie di appartenenza;
- aiutare a risolvere (almeno parzialmente) il problema della sinonimia;
- monitorare i *tag* più utilizzati e la loro correlazione con le categorie, per correggere o integrare la tassonomia (il sistema di classificazione creato dall'alto);
- assorbire progressivamente in un sistema di metadati o in un vocabolario controllato i *tag* che superino una certa soglia d'uso - in modo simile a quanto avviene nel sito della BBC (Loasby, 2006).

## Bibliografia

- Bates, M., *The design of browsing and berrypicking techniques for the online search interface*. "Online Review", 13 (October 1989), p. 407-424. Disponibile anche alla URL <<http://www.gseis.ucla.edu/faculty/bates/berrypicking.html>>
- Bowker, G. C. - Star, S. L., *Sorting things out. Classification and its consequences*. MIT Press, 1999
- Carcillo, F. - Rosati, L., *Tags for citizens. Integrating top-down and bottom-up classification in the Turin municipality website*, in Schuler Douglas (ed.), *Online communities and social computing, Second International Conference, OCSC 2007*. Held as part of HCI International 2007 (Beijing, China, July 22-27, 2007).

- Proceedings Series: "Lecture Notes in Computer Science", Vol. 4564 (2007), Springer, p. 256-264
- Gnoli, C. - Marino, V. - Rosati, L., *Organizzare la conoscenza: dalle biblioteche all'architettura dell'informazione per il Web*. Milano : Tecniche Nuove, 2006
- Lakoff, G., *Women, fire and dangerous things. What categories reveal about mind*. University of Chicago Press, 1987
- Loasby, K., *Changing approaches to metadata at bbc.co.uk: from chaos to control and then letting go again*. "ASIS&T Bulletin", October/November 2006  
<<http://www.asis.org/Bulletin/Oct-06/loasby.html>>
- Pirolli, P., *Information foraging theory. Adaptive interaction with information*. New York : Oxford University Press, 2007
- Pirolli, P. - Card, S., *Information foraging*. "Psychological Review" 106 (1999), n. 4, p. 643-675
- Quintarelli, E., *Folksonomies: power to the people*, in *Proceedings of ISKO Italy and University of Milano Bicocca Meeting (Milan, 2005)*. <<http://www.iskoi.org/doc/folksonomies.htm>>
- Rosati, L., *I due assi dell'architettura dell'informazione*. "Trovabile", 2006  
<<http://trovabile.org/articoli/assi-architettura-informazione>>
- UK Cabinet Office eGovernment Unit, *IPSV - Integrated Public Sector Vocabulary*. 2006 <<http://www.esd.org.uk/standards/ipsv/>>.

