

Accesso all'informazione per codificazione o per condivisione ? Le "folksonomie".

**“Interoperabilità delle biblioteche digitali:
costituzione, diffusione e condivisione delle risorse informative”**

Roma, 10 ottobre 2006

Nicola Benvenuti

(Università di Firenze)

Social tagging e biblioteche digitali

- Il social tagging corrisponde almeno a quattro caratteristiche espresse nel “Manifesto delle biblioteche digitali”
 - 1. *Le biblioteche digitali sono conversazioni* Non biblioteca digitale, ma biblioteche digitali, non un sistema, una grande narrazione sistematica, ma tante conversazioni tenute insieme da un linguaggio comune, da una struttura comunicativa basata sull'assunzione di impegni fra comunità diverse per pubblici diversi.
 - 5. *Le biblioteche digitali non si occupano di techeology* Spesso le conversazioni sulle biblioteche digitali sono condizionate da posizioni dove gli strumenti si presentano come un misto di tecnologia e di ideologia (techeology) e diventano essi stessi fini e non mezzi per offrire servizi.
 - 8. *Le biblioteche digitali condividono i problemi delle loro comunità* Le biblioteche digitali basano il proprio rapporto con gli utenti sul metodo dell'interazione, favoriscono la partecipazione attiva degli utenti alla propria comunità e cercano di parlare il loro linguaggio.
 - 10. *Le biblioteche digitali sono in rete.* Tutti i soggetti che hanno accesso ad Internet (ovvero alla "rete") sono potenziali fruitori delle risorse digitali

Folksonomy

- Uso di termini scelti dagli utenti e tratti dal linguaggio per designare oggetti informativi testuali e di altri media (es. foto, movies, musica ...)
- La scelta di parole chiave da parte degli utenti non è una novità:
 - Parole chiave assegnate dagli autori ai propri articoli scientifici (es medline, o anche gli archivi OAI)
 - Strutture gerarchiche create dagli utenti (il primo yahoo e successivamente DMOZ)
 - Correzioni in CiteSeer
 - Metadati creati dagli utenti
 - (...)
- Da questa attività di tagging può derivare la costruzione di un vocabolario vicino alle esigenze informative degli utenti in grado di migliorare il recupero dell'informazione e di abbattere I costi di categorizzazione delle risorse web tramite specialisti ?

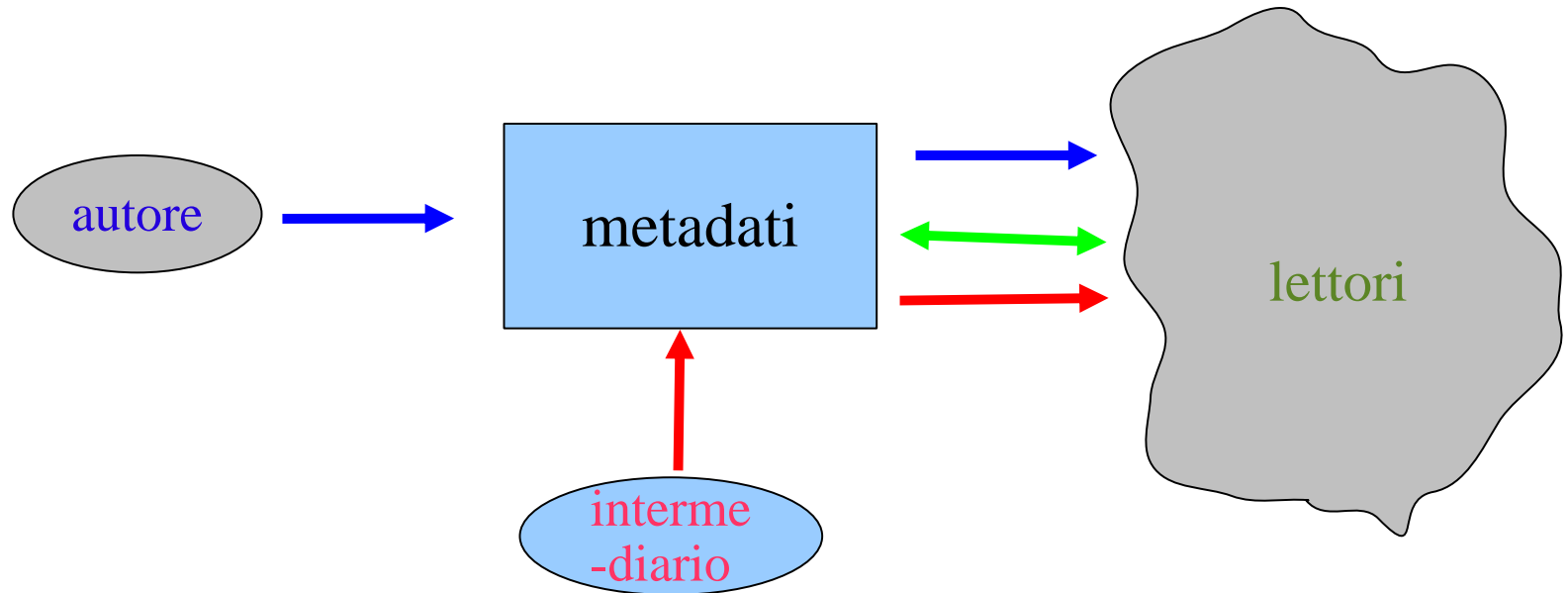
Social tagging

- “Creazione di metadati degli utenti condivisi in una comunità” (A. Mathes)
 - Accento sui metadati (automatici, manuali)
 - Comunità come dimensione di sviluppo (iscrizione, target, finalità ...)
- Il tagging è accompagnato da una serie di altre funzionalità per favorire interazioni cooperative
 - commenti,
 - raccomandazioni di utenti con analoghi interessi,
 - related tag,
 - Possibilità di costruire gruppi omogenei
 - ...

Disintermediazione

- Impossibilità di “catalogare” la rete utilizzando esperti classificatori
- ...non solo necessità:
 - ambiente di comunicazione di rete e di protagonismo degli utenti del web 2.0: creare non solo consumare; scrivere non solo leggere (Tim O’Reilly: Architettura della partecipazione)
 - Limiti delle classificazioni bibliotecarie

Relazioni tra gli attori della comunicazione



Catalogazione di esperti (formati catalografici, authority files ...)

Catalogazione degli autori (es. DC; parole chiave in articoli; Flickr ...)

Folksonomy

Personal, narrow e broad volksonomy

- Si distinguono almeno tre funzioni del tagging:
 - *Personal Volksonomy*: consiste nell'assegnare un tag ad una risorsa per uso personale, cioè in ambiente chiuso (es. gmail)
 - *Narrow volksonomy*: consiste nell'utilizzo dei tag per dare accesso con il linguaggio a risorse non testuali, come foto o film (es. flickr); in genere i tag sono pochi e assegnati dall'autore della risorsa
 - *Broad volksonomy*: serve a organizzare le risorse in ontologie condivise (es. del.icio.us, connotea ...) e consiste spesso di molti termini assegnati ad una stessa risorsa
- più tag sono assegnati ad una risorsa, più è assicurata la reperibilità del documento.

Classificazioni e volksonomies

http://www.shirky.com/writings/ontology_overrated.html

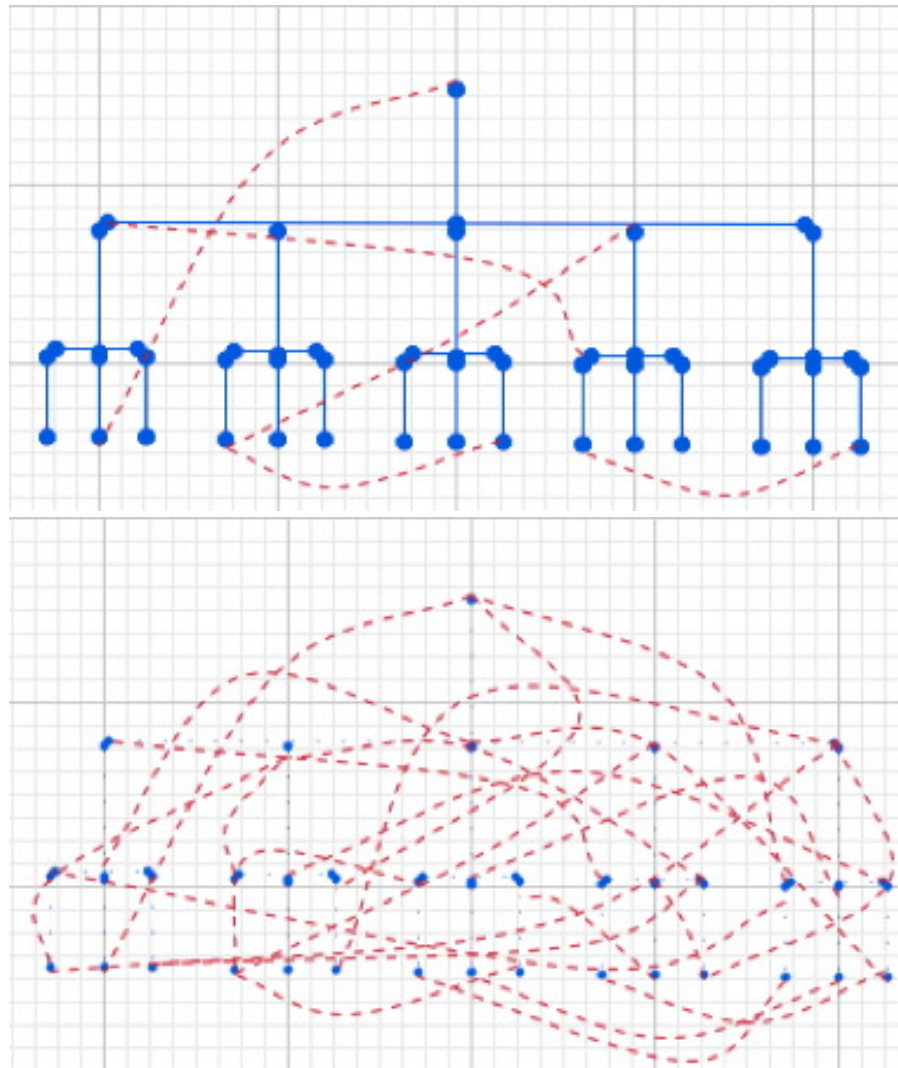
- La categorizzazione degli utenti non rappresenta una estensione bensì una *rottura radicale* con le precedenti strategie di categorizzazione (Clay Shirky)
 - Le classificazioni gerarchiche non rappresentano la realtà bensì sono costruzioni parziali e ideologiche
 - L'utilità di classificazioni gerarchiche nasce dai vincoli fisici derivanti dal posizionamento dei libri sugli scaffali
 - In realtà nelle classificazioni si fa confusione tra libro (*contenitore*) e concetti (*contenuto*)
 - Internet non ha vincoli fisici
 - Il tagging, cioè url+etichetta, rinvia ai *significati*.

...continua

http://www.shirky.com/writings/ontology_overrated.html

- Orientamento al mercato: emergono “*motivazioni individuali*, ma *valori di gruppo*”. Importante non è “chi categorizza meglio di me”, bensì “chi categorizza come me”.
- È importante sapere *chi* ha fatto un certo tag ad una risorsa (cioè quale profilo di interessi ha) e *quando* (quanto è aggiornato il tag)
- Il filtro di qualità viene *dopo* la pubblicazione ed è costituito dai link che riceve
- Il collegamento tra schemi individuali non avviene a livello di tag uguali, bensì tramite i documenti conosciuti e identificati tramite URL (identificatori univoci): all’inizio vi è il contenuto, da esso si giunge ai punti di vista (tag)
- La corrispondenza tra tag (es. sinonimia) non è binaria (si/no) ma *analogica*, cioè prefigura sovrapposizione parziale di domini e quindi implica diversità di contenuto

Struttura gerarchica e piatta



Tag inseriti dagli utenti: pro/contro

Contro

- Ambiguità dei termini usati
 - Polisemia, singolare/plurale, grafia, lingue, ...
 - Tag composti da una (del.icio.us, Flickr) o da più parole (connotea)
 - Non c'è gestione dei sinonimi
- Assenza di contestualizzazione disciplinare
- Nessuna uniformità di uso

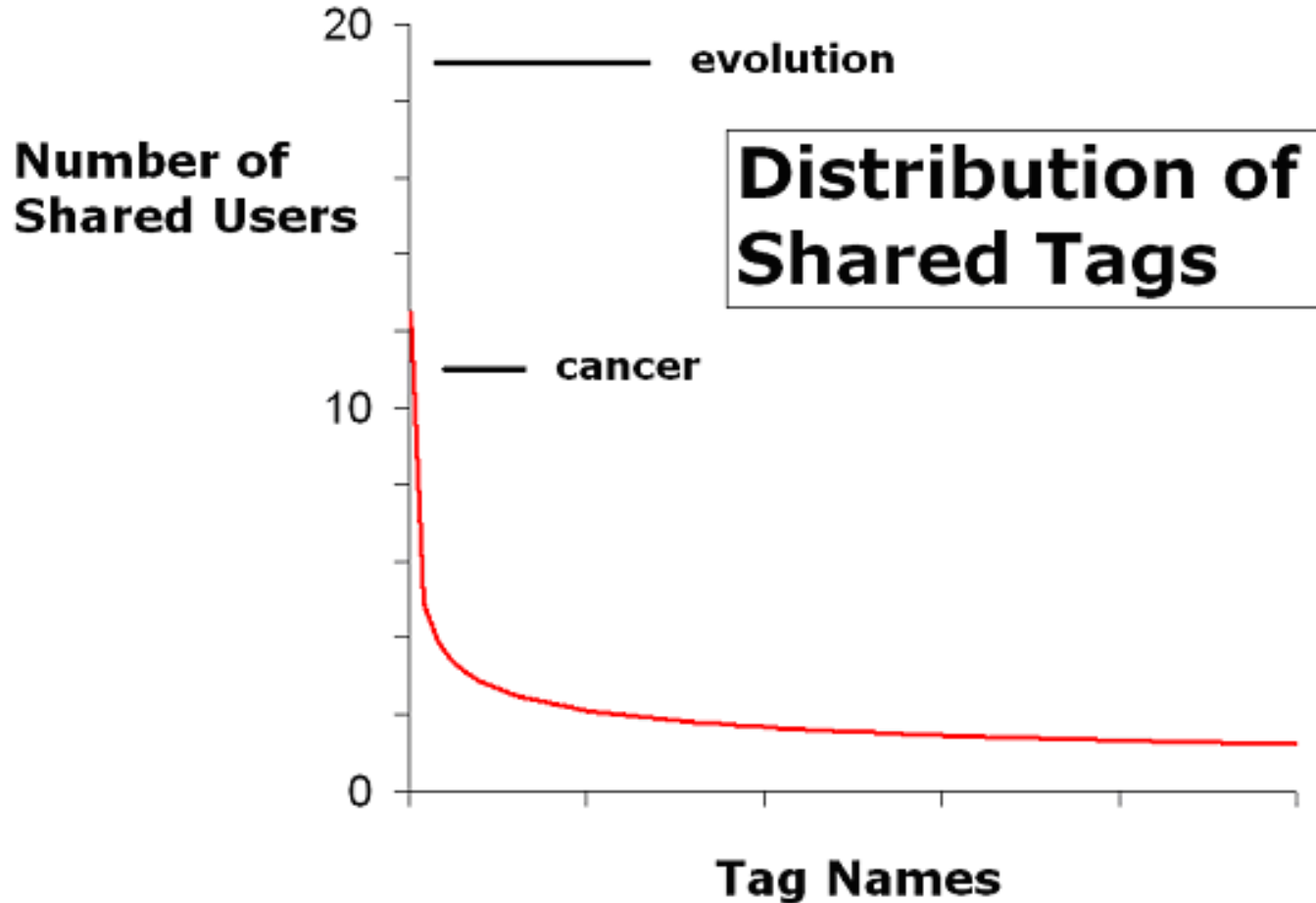
Pro

- Serendipity
- Trovare contenuto interessante (*discovery*) – *contra* – ricerca diretta di documenti rilevanti secondo una query (*find*)
- Possibilità di feedback immediato (J. Udell), cioè di visualizzare il dominio semantico dei tag (... e attivare una negoziazione implicita)
- Abbassa le barriere della cooperazione: assenso su un *significato* non su una *gerarchia*
- Impulso a condividere
- *Desired lines* (P. Merholz): utilizzare i tag più popolari per identificare un vocabolario controllato più vicino al linguaggio degli utenti

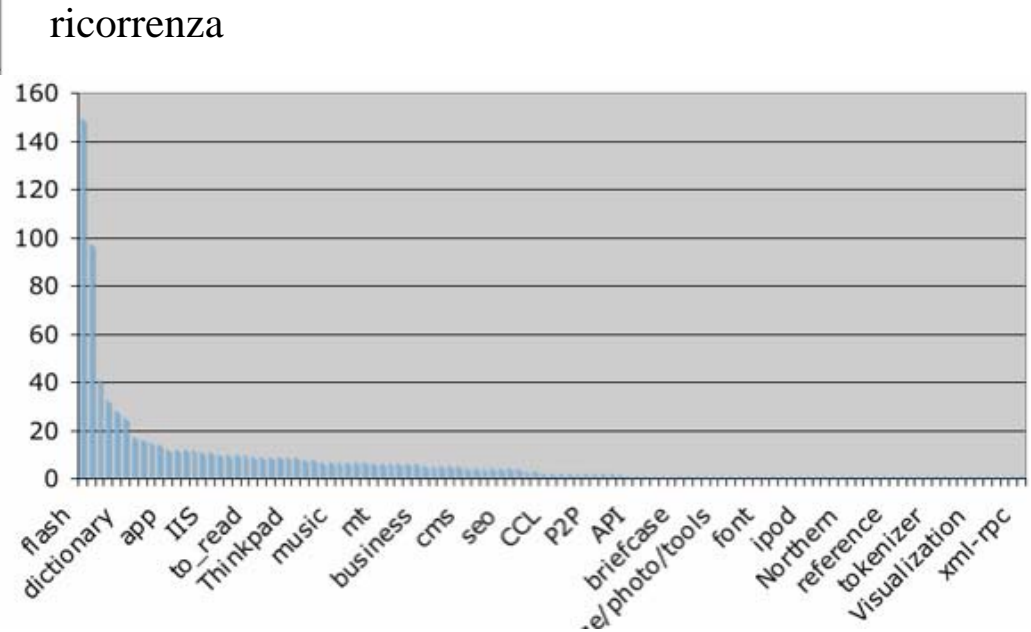
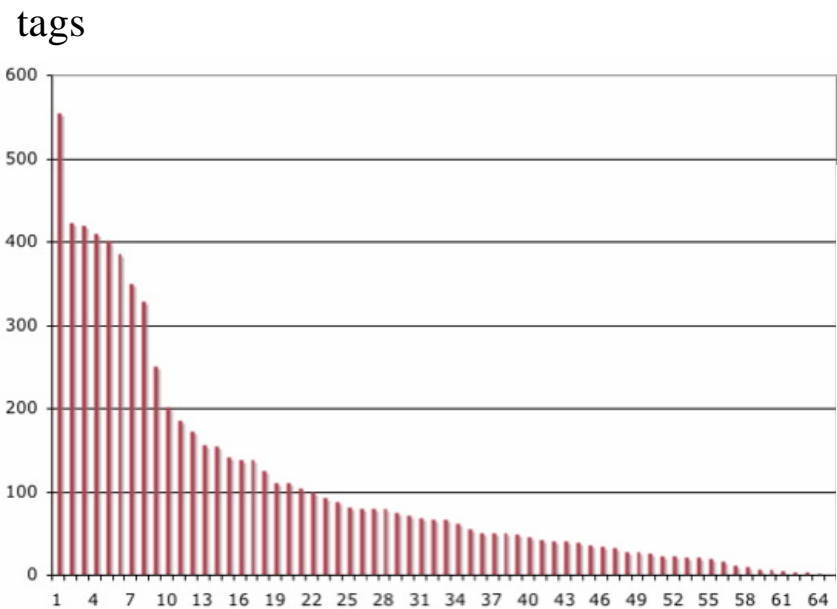
La proposta di Merholz: seguire le “linee desiderate”



Uso dei tag



Tags in del.icio.us (C. Shirky)



Findability

- Tag cloud:
 - del.icio.us <http://del.icio.us/tag/>
- Faccette:
 - Fac.etio.us
<http://demo.siderean.com/facetious/facetious.jsp>
 - Rawsugar <http://blog.rawsugar.com/bsamples/eng-tb3-r.html>
 - Mefedia <http://mefedia.com/tags/>
- Clustering
 - [grokker](#)

Comunità di rete

- È formata da individui che interagiscono
- Vi è uno scopo comune
- Sono presenti politiche condivise (presupposti taciti, rituali, protocolli, regole e leggi che guidano l'interazione tra gli individui)
- Sistemi informatici per supportare le interazioni sociali e trasmettere senso di appartenenza

E-commerce

- Il primo a utilizzare i tag per l'autorganizzazione (individuale) delle informazioni è stato google in gmail
- Amazon
 - *tag this product* e' inteso:
 - a) come modo per trovare in amazon libri dello stesso contenuto e per ricordarseli facilmente. La prima funzione implica condivisione.
 - b) per organizzare in modo personale i titoli comprati
 - c) ricordare I libri da comprare (per spesa, per marca o modello)
 - d) selezionare regali col nome del destinatario per poi fare una scelta
 - e) sulla base dei tag vengono proposte raccomandazioni personalizzate (recommendations by your tags)
 - f) il sistema tiene memoria dei propri tag per estendere la ricerca
 - g) Solo se si compra qualcosa i propri tag diventano pubblici e si possono vedere i tag degli altri, (organizzati collettivamente dai clienti di amazon)
 - h) i tag possono anche essere privati, editati o rimossi
 - La condivisione si attua dopo che si diventa clienti.

Connotea

- Creata dal Nature Publishing Group è un software di Social Bookmarking rivolto ad un ambiente accademico
 - Utilizzo di script per il browser per estrarre dati e metadati strutturati
 - Possibilità di configurare un resolver OpenUrl per accedere al full text nelle raccolte della propria istituzione
 - Possibilità aggiungere descrizione, tag e commenti
 - Possibilità di creare gruppi di utenti
 - Possibilità di dichiarare ogni bookmark “condiviso” “condiviso nel gruppo” o “privato”
 - Indicazione di chi altri ha archiviato una certa risorsa e come l’ha categorizzata
 - Operazioni con i tag:
 - lista dei tag utilizzati da ogni utente
 - Editing dei tag
 - possibilità di creare “scope note”
 - Suggerimento di tag correlati
 - Indicazione di utenti che hanno archiviato le stesse risorse o usano gli stessi tag

Comunità come learning space

- Conoscenza tacita (es. abilità di fare ricerche su internet per trovare documenti rilevanti)
- Conoscenza connettiva (D. De Kerckhove) tramite le relazioni tra utenti e le connessioni tra i tag che permette di condividere, confrontare, selezionare le conoscenze
- L'informazione non viene consumata passivamente ma viene valorizzata nella condivisione, nella interazione con gli altri, in modo sia esplicito che implicito.

Classificazioni e folksonomies

- | | |
|---|---|
| 1) Costosa da mantenere | 1) Non richiede manutenzione |
| 2) E' difficile da applicare | 2) Non richiede preparazione specifica |
| 3) E' difficile da capire per l'utente | 3) E' di immediata comprensione per l'utente |
| 4) E' rigida e non aggiornata tempestivamente | 4) E' aggiornata in tempo reale |
| 5) preordinata | 5) postordinata |
| 6) Multilivello | 6) Piatta |
| 7) Contestualizza i concetti tramite gerarchie disciplinari | 7) Contestualizza i concetti a livello di significati |
| 8) Moltiplicazione tramite codificazione | 8) Moltiplicazione tramite condivisione |

Moltiplicazione

Enzo Rullani, *La fabbrica dell'immateriale: produrre valore con la conoscenza*, Carocci, 2004

- Il valore della informazione/conoscenza in termini sociali consiste anche nella possibilità di *moltiplicazione*, cioè nel numero di riusi che può avere
- *Condivisione* e *codificazione* sono metodologie diverse di moltiplicazione della conoscenza che vengono intese sempre meno in contrapposizione. Hanno effetti ed efficacia diversi ma, dati i differenti costi di implementazione, la loro applicazione dipende dall'ampiezza dei contesti in cui operano
 - La *condivisione* è un sistema molto efficiente di organizzazione delle conoscenze, che nasce dalla partecipazione e progettazione di una medesima esperienza, ma genera un ridotto bacino d'uso (condivisione familiare, locale, professionale ...). Ampliare il bacino può avere costi eccessivi
 - La *codificazione* mira invece a definire ex ante e dall'alto gli standard a cui contesti ed esperienze devono uniformarsi. L'allineamento delle diverse visioni è delegato a specialisti che definiscono le condizioni necessarie per raggiungere l'effetto voluto. Dopo di che gli attori devono adottare, anche separatamente, il punto di vista standard. La codificazione è dunque adatta a processi di propagazione che puntano ai grandi numeri e all'uso di conoscenza standardizzata

Comporre l'approccio della condivisione con quello della codificazione

- OCLC rende WorldCAT ricercabile da Google per far sì che il catalogo acquisti valore tramite internet
- Anche per trovare informazione con strumenti semantici è possibile un analogo approccio utilizzando le *folksonomie* create in ambiente locale (territoriale/disciplinare) per accedere a informazioni bibliografiche codificate in modo gerarchico e con linguaggi controllati
 - Possibilità di scoprire *significati* non segnalati dalla classificazione del contenuto di libri in catalogo
 - Possibilità di ritrovare *tipologie di risorse diverse* ad esempio collegando i tag attribuiti in connesione con quelli assegnati ai libri del catalogo...
 - Possibilità di offrire un *accesso flessibile* alle classificazioni gerarchiche, diminuendo fortemente la curva dell'apprendimento delle regole di codificazione.