

Accesso alla conoscenza nascosta

di Nicola Benvenuti

Il compito delle biblioteche consiste nel raccogliere, gestire e mettere a disposizione degli utenti, nel modo più funzionale ed efficace ai fini del reperimento dell'informazione, i risultati del processo di creazione della conoscenza formalizzati in pubblicazioni. In questo senso gli strumenti principe delle biblioteche sono i cataloghi delle monografie e delle riviste e le banche dati che contengono lo spoglio sistematico delle riviste scientifiche e accademiche.

A ben vedere però essi, pur rimanendo essenziali e insostituibili, non esauriscono più, nell'era di Internet, la gamma degli strumenti di accesso all'informazione. Ciò è dovuto non tanto a motivi di tipo tecnologico, anche se alla base di questo sviluppo vi è sicuramente la facilità di pubblicazione elettronica, quando ai profondi mutamenti indotti nel ciclo della comunicazione scientifica dall'evoluzione dei modi di creazione e organizzazione della conoscenza. In tal modo emergono vasti settori di conoscenza che per difficoltà di codifica e di valorizzazione o per il fatto di aver svolto fino a ieri un ruolo secondario, sono definibili dal punto di vista dell'accesso come conoscenza "nascosta".

Un buon punto di partenza per una riflessione su questi temi è l'analisi del ciclo delle pubblicazioni scientifiche, elaborato da Krishna Subramanyam all'inizio degli anni Settanta, che riconnette le diverse fasi della creazione della conoscenza e della comunicazione scientifica¹. Il modello, per quanto semplificato, ben esprime ancora la situazione odierna², e ulteriori considerazioni significative emergono quando si riconnette il ciclo della creazione di conoscenza oltre che al ciclo della pubblicazione, al ciclo dell'accesso³, più consono all'area di intervento delle biblioteche: è qui che si evidenziano almeno tre aree problematiche.

¹ Maurizio Zani, Brian C. Vickery, *Scientific communication in history*, "Bibliotime", V, 3, novembre 2002; <<http://didattica.spbo.unibo.it/bibliotime/num-v-3/zani2.htm>>.

² Sheila R. Curl, *Subramanyam Revisited: Creating a New Model for Information Literacy Instruction*, in "College and Research Libraries", 62, 5 settembre 2001 (7.01.2006); <<http://www.ala.org/ala/acrl/acrlpubs/crljournal/backissues2001b/september01/curl.pdf>>.

³ Si veda la rielaborazione grafica del modello a cura di due bibliotecarie della Washington University, Carol Green e Patty Carey (7.01.2006); <<http://www.lib.washington.edu/subject/environment/imt220/pubcycle.jpg>>.

LA LETTERATURA GRIGIA

La prima riguarda quella galassia della produzione scientifica che precede l'offerta dei risultati di una ricerca al *peer review*, cioè al comitato di redazione di una rivista scientifica. Tale attività o non si concretizza in pubblicazioni formali oppure dà luogo alla cosiddetta "letteratura grigia", un insieme di manifestazioni che nel ciclo della pubblicazione compaiono come scarsamente distribuite e quindi non accessibili affatto o solo in ambiti e modalità limitati. La facilità di pubblicare in formato elettronico e di distribuire qualsiasi documento in rete ha portato in primo piano questi materiali, ne ha sollecitato la produzione e ne ha evidenziato anche il valore scientifico. Infatti essa – diversamente dalla tradizione bibliografica che definisce letteratura grigia i testi a stampa con meno di 99 pagine – è costituita da tipologie di documenti che possono avere diverse caratteristiche: essere espressione dell'attività di enti, istituzioni e individui (relazioni sulle attività di enti pubblici e privati – p. e. *report* o cataloghi di prodotti –, resoconti di convegni e incontri, tesi, elaborati individuali...); documenti che richiedono per la loro stessa natura, continui aggiornamenti e nuove versioni (news, rapporti periodici, blogs...); documenti che hanno anche un corrispettivo a stampa ma che gli autori depositano *online* nella consapevolezza che in rete la loro visibilità e quindi la loro circolazione è assai più ampia. Questa informazione si manifesta in gran parte nella struttura tipicamente *bottom-up* della rete e lungo *link* e *bookmark* che connettono a siti web.

Il modello evidenzia inoltre che, nelle prime fasi delle ricerche, si produce una conoscenza non destinata alla pubblicazione a stampa, che potremmo anzi definire preparatoria o anche residuale rispetto alla definitiva formalizzazione dei risultati della ricerca, cioè appunti e annotazioni varie, carteggi, progetti di ricerca, schemi esemplificativi, modelli, tracce di interventi e di corsi, set di dati, elaborati a vari livelli, raccolte di *link*, etc.

Oltre che per varietà tipologica, ciò che caratterizza questa letteratura è l'essere svincolata dal meccanismo consolidato di certificazione scientifica, quello cioè che passa attraverso la via principe della valutazione di comitati scientifici, ma non ne è affatto priva, o non del tutto: tale certificazione può essere infatti conferita o dalla natura dell'ente che ospita i documenti, oppure più spesso può nascere dai meccanismi stessi di comunicazione, organizzazione e collaborazione nell'ambito della rete: in una parola è conferita dal controllo esercitato dalle comunità di rete stesse, naturalmente ai diversi livelli di competenza, tramite *link*, citazioni, segnalazioni...⁴

FRAMMENTARIETÀ DELL'INFORMAZIONE

La seconda area problematica nasce dalla considerazione delle modalità di accesso alle pubblicazioni in un ambito di rete. Le innovazioni di maggior rilievo in questo

⁴ Il tema oggetto di discussione nell'ambito della filosofia della conoscenza e ha trovato particolare attenzione nel dibattito sullo e-learning. Si veda ad esempio: Jenny Preece, *Comunità on line: progettare l'usabilità, promuovere la socialità*, Milano, Tecniche nuove, 2001.

ambito sono state sicuramente il trasferimento su supporto elettronico prima e la migrazione in rete poi, delle banche dati bibliografiche, soprattutto quelle contenenti spogli di articoli di riviste, e delle banche dati di *full text*. Queste cruciali innovazioni stentano però a modificare come auspicabile il flusso dell'informazione riconnettendosi in procedure coerenti e integrate e creando nuovi i modelli di servizi corrispondenti alle attese degli utenti sempre più condizionati dall'esperienza della rete. Proprio dalla rete nasce infatti un modello di servizi che realizza in un *continuum* le funzioni di ricerca dell'informazione, di localizzazione e di uso immediato.

Questo *continuum* appare ancora spezzato nell'attuale sistema dei servizi bibliotecari. All'enorme potenziamento delle capacità di recupero dell'informazione e di disponibilità di *full text* si contrappone una frammentazione di soggetti, produttori, distributori, editori, per non parlare delle risorse disponibili in rete, ognuno con propri formati dell'informazione, con propri linguaggi di ricerca, con proprie piattaforme software, con proprie modalità di accesso, con propri criteri di assemblaggio dei dati, etc. L'informazione appare strutturata in *silos* non comunicanti, non ispirati a standard e con struttura e gestione proprietaria (il "web nascosto" ma anche l'ambiente informativo della biblioteca).

CONOSCENZA *TACITA* E CONOSCENZA *COLLETTIVA*

Il modello di Subramanyam ci segnala, infine, che parte della creazione e trasmissione della conoscenza si attua attraverso il cosiddetto *invisibile collegé*⁵, con cui oggi si intende la trasmissione di conoscenza attraverso la comunicazione interpersonale e la collaborazione in comunità coese, esternando e condividendo la conoscenza implicita di individui e gruppi...

In senso più generale l'affermazione di Polany: "Sappiamo più di quello che riusciamo a dire"⁶ rinvia alla esistenza accanto ad una cultura esplicita e formale, di una conoscenza implicita di cui possiamo anche non avere consapevolezza e che costituisce il nostro "capitale intangibile". Si tratta di una conoscenza individuale, per lo più relativa al saper fare, il *know how*, contrapposto alla conoscenza razionale, il *know about*.

Accanto ad essa gli studi di filosofia della conoscenza sottolineano l'esistenza di una conoscenza collettiva: "La formazione di una 'conoscenza comune' il bisogno di stabilire regole e linguaggi per facilitare la formazione collettiva della conoscenza e il riconoscimento che il 'contesto' ha un peso nel processo di creazione della conoscenza,

⁵ L'espressione richiama storicamente il cambiamento epistemologico radicale della modernità, dalla conoscenza come rivelazione alla conoscenza come produzione umana tramite l'esperienza; mi piace qui rinviare alla voce corrispondente in Wikipedia; <<http://www.wikipedia.org>>.

⁶ Michael Polanyi, *Conoscenza personale: verso una filosofia post-critica*, Milano, Rusconi, 1990 (trad. it. di: *Personal knowledge*).

sono caratteristiche che emergono in modo naturale e sottolineano che la conoscenza risulta da un processo di costruzione collettiva⁷.

Internet ha letteralmente fatto esplodere queste forme di conoscenza offrendo l'ambiente ideale per la loro manifestazione e trasmissione. Un esempio è costituito da Google stesso, cioè da un motore di ricerca che affida la rilevanza dei risultati reperiti al criterio della "popolarità" (ovviamente, solo in parte sinonimo di comunità o di gruppi di interesse) sfruttando i metadati impliciti creati dai comportamenti degli utenti della rete tramite i *link* ai siti più valutati. In tal senso delinea una organizzazione dell'informazione in rete radicalmente diversa da quella dei cataloghi bibliografici e compendiate nei *subject gateway*, ma anche in Yahoo o l'Open Information Project, e basato sulla mediazione tra autore e utente tramite surrogati dei documenti creati da personale specializzato ed espressi in forma strutturata e normalizzata. La metafora di Google risulta invece basata su parole del linguaggio naturale e sulla articolazione comunicativa delle comunità di interesse: una metafora quindi che, riprendendo il suggerimento del punto 1 del "Cluetrain Manifesto" potremmo chiamare delle "conversazioni"⁸. A questi diversi approcci corrispondono in parte anche due tipologie diverse di conoscenza: la prima, racchiusa nelle pubblicazioni formali, libri e articoli, secondo strutture definite e provate (presentazione del problema, analisi dei dati, dimostrazione, conclusione), è esplicita e codificata; la seconda, costituita da "comportamenti" collaborativi (le comunità), capacità di muoversi in rete ("saper fare"), di comunicare con gli altri (esternalizzare e interiorizzare conoscenze implicite), di cogliere e costruire nuove connessioni tra eventi (si pensi al meccanismo della *serendipity* nella ricerca di rete) è sostanzialmente tacita.

IL KNOWLEDGE MANAGEMENT ALLA PROVA

Queste aree problematiche, che nascono dall'analisi delle modalità di accesso alle diverse fasi di creazione della conoscenza, non possono non investire e trasformare il compito delle biblioteche nel mediare l'accesso all'informazione, rendendolo più complesso: non solo perché emergono nuove tipologie di documenti e si affermano nuovi modelli di servizi di informazione, ma perché accedere alla conoscenza significa anche creare il contesto adeguato per esternalizzarla, e condividerla.

Per esaminare questi aree di creazione di conoscenza al fine di permetterne l'accesso, i tradizionali strumenti della biblioteconomia non sono sufficienti ed è necessario testare nuovi approcci come quelli forniti dalle discipline del *knowledge management* (da ora in poi KM).

⁷ Bernard Ancori, Antoine Bureth and Patrick Cohendet, *The economics of knowledge: the debate about codification and tacit knowledge*, in "Industrial and corporate change", 9, 2, 2000, pp. 276.

⁸ Si veda la traduzione italiana del "Cluetrain Manifesto": 95 TESI; <http://www.metierediscrivere.com/testi/Tesi.htm>. <http://www.economist.com/displaystory.cfm?story_id=4484119>.

Il KM è disciplina relativamente recente e continua a rivestire un ruolo di primo piano nella interpretazione delle profonde modificazioni create dallo sviluppo della società della conoscenza. Non a caso anche in recenti convegni di natura biblioteconomica cominciano a comparire sezioni dedicate al KM, vedi nel 2005 la “34th Annual Conference” di Liber a Groningen o la stessa Conferenza IFLA di Oslo⁹. D'altra parte il KM nella sua evoluzione si misura via via con discipline e metodologie contigue a quelle biblioteconomiche, quali il *content management* o le ontologie, per non citarne che alcune.

In un articolo del 2000 Elisabeth Cronin e Blaise Davenport, due autori di primo piano del KM hanno delineato tre approcci all'organizzazione della conoscenza in ambiente accademico, discutendone l'efficacia dal punto di vista delle premesse e degli obiettivi del KM stesso. Nel riprendere questa suddivisione va avvertito che l'opinione degli autori propende nettamente per considerare campo specifico del KM la terza interpretazione proposta, l'esternalizzazione della conoscenza implicita, e anche che altri criteri sono stati suggeriti per caratterizzare le peculiarità di questa disciplina. Tuttavia questo approccio ci sembra utile per illustrare alcuni processi che coinvolgono le biblioteche, con la raccomandazione di non intenderli in alcun modo come successivi o nettamente separati, ma anzi integrati e compresenti. Spero infatti di riuscire a evidenziare che essi si manifestano variamente in tutte le esemplificazioni che saranno proposte.

KM1

Il primo approccio è quello della *Library Information Science* (LIS) ed è incentrato sulla organizzazione di risorse interne ed esterne all'impresa o all'ente, ai fini di archiviazione, distribuzione e conservazione. Il focus, se torniamo alla rappresentazione della comunicazione scientifica delineata sopra, riguarda proprio la nebulosa delle nuove tipologie di documenti (testuali o multimediali) che hanno reso la ricerca in rete strumento indispensabile anche in ambienti scientifici.

Ponendo l'accento sulla tipologia dei materiali questo approccio porta ad evidenziare soprattutto due nuove tipologie di collezioni: la costruzione presso le biblioteche di archivi OAI; la catalogazione di siti internet nei *subject gateways*.

OAI: I DATA PROVIDER

Gli archivi aperti sono in genere presentati come archivi per il deposito di una precisa tipologia di documenti, i *preprint* e i *postprint* di articoli delle riviste scientifiche.

⁹ Gli atti sono consultabili all'indirizzo: <http://www.ensib.fr/article.php?id=206&cat=Biblioth%C3%A9que+num%E9rique&id_cat=206>; per World Library and Information Congress: 71st IFLA General Conference and Council; <<http://www.ifla.org/IV/ifla71/Programme.htm#13August>>.

In realtà il ruolo degli archivi aperti non si esaurisce in questa funzione pur essenziale di memorizzare e preservare nel tempo articoli pubblicati sulle riviste scientifiche che le istituzioni di ricerca ed educative perseguono al fine di non disperdere gli ingenti investimenti finanziari messi in atto per creare conoscenza. A partire dal progetto *Dspace*, infatti, si è evidenziata una ulteriore funzione, quella di archiviare e organizzare quella documentazione che si costruisce ai margini dell'attività didattica e di ricerca e precipita in una vasta tipologia di documenti fatta di set di dati, appunti, annotazioni, versioni etc. Riferendosi all'ambito universitario è stato osservato:

La comunità accademica ha raccolto informazione tacita per anni. Nota come annotazioni a margine e appunti, quest'informazione ha spesso preso la forma di *bookmark*, fisici o elettronici. (...) Ma per la maggior parte l'informazione tacita ha preso la forma di "strumenti operativi" (*tools n.d.t.*) o "di assistenza" (*aids n.d.t.*) creati da individui. Questa informazione non è sistematizzata e sicuramente non è a disposizione della più vasta comunità. Perché pensare di svincolare queste colossali fonti di informazione così difficili da gestire? La risposta sta nel potenziale di crescente produttività e innovazione che potrebbe ottenersi se questa informazione potesse in qualche modo essere integrata nei processi della ricerca scientifica¹⁰.

Dspace è un superarchivio lanciato nel settembre 2002 (ma progettato già nella primavera del 2000) dalle biblioteche del MIT in collaborazione con Hewlett-Packard. Esso nasce proprio per l'esigenza di conservare e preservare (la D rinvia ai termini "Documenti Digitali Durevoli") materiali accumulatisi nei computer degli studiosi:

Ogni anno i ricercatori dello MIT creano almeno 10.000 testi, file di dati, immagini, collezioni di annotazioni, oltre ad audio e video clips. La ricerca trova spesso destinazione nelle riviste scientifiche ma il materiale residuale rimane rintanato sui personal computer, i siti web e i server dei dipartimenti. È accessibile solo a pochi ancor oggi. E con la rapida evoluzione di computer e software si avvicina il momento in cui i files salvati oggi non saranno più accessibili per nessuno¹¹.

Nel presentare lo scopo del progetto, MacKenzie Smith¹², allora direttore associato alla tecnologia dello MIT, citava insieme ai *preprint* e agli articoli: relazioni tecniche, documenti di lavoro corrente, interventi a conferenze, tesi, set di dati, matlab, immagini, file audio, oggetti didattici (*learning object*), collezioni bibliografiche digitali riformattate. La dimensione organizzativa è rappresentata naturalmente dalla comunità sia in senso istituzionale (Dipartimenti, Facoltà ...) che disciplinare.

¹⁰ Brian Hawkins, *Libraries, Knowledge Management, and Higher Education in an Electronic Environment*, "Capitalising on knowledge the information profession in the 21st century", 24-26 October 2000 (ALIA 2000); <<http://conferences.alia.org.au/alia2000/proceedings/brian.hawkins.html>>.

¹¹ Sally Atwood, *MIT's Superarchive*, in "Technology review", December 2002/ January 2003; <<http://www.technologyreview.com/articles/02/12/atwood1202.asp?p=1>>.

¹² Dspace [presentazione ppt] / MacKenzie Smith; <<http://www.dpconline.org/graphics/events/presentations/pdf/DspaceatDPCOct2002.pdf>>.

DSpace costituisce pertanto il tentativo di inserire organicamente questa conoscenza non valorizzata nel flusso dell'informazione accademica ed è perciò destinato a cambiare profondamente il modo in cui il mondo universitario raccoglie e dissemina i risultati delle ricerche.

OAI: I SERVICE PROVIDER

Come è noto nell'ambito OAI i soggetti attivi si suddividono oltre che in *data provider* – in pratica chi raccoglie i dati - anche in *service provider* – cioè chi crea servizi su i dati raccolti. Nell'ambito dei *service provider* la funzione sicuramente più sviluppata è costituita dallo *harvesting* dai vari archivi che espongono i metadati secondo il protocollo OAI-PMH. *Oaister*¹³ con 5,914,431 record provenienti da 536 istituzioni (al 30 settembre 2005) è certamente l'*harvester* che raccoglie più dati, ma ne esistono numerosi altri come in Italia PLEIADI (Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali)¹⁴ che scaturisce dalla collaborazione fra due importanti consorzi interuniversitari italiani, CASPUR e CILEA, nell'ambito del progetto AEPIC.

Oltre allo *harvesting* sono stati studiati altri servizi individuali e collettivi volti ad esaltare la funzione degli archivi OAI per facilitare la ricerca individuale e soprattutto come luoghi di collaborazione. Il programma europeo "Cyclades"¹⁵, svoltosi tra il 2000 e il 2003, si riproponeva proprio il compito di riarticolare in questo senso il ruolo del *service provider*. Utilizzando tecniche di *Basic Support for Collaborative Work* (BSCW) sviluppate dal Fraunhofer Institute for Applied Information Technology (FIT), il progetto disegnava una struttura collaborativa basata su *folder* individuali e di comunità per raccogliere documenti di diversa tipologia, testi, *link*, *queries*, forum di discussione, etc. Le funzionalità previste andavano dalla possibilità di assegnare un voto ad ogni documento (che in tal modo riceveva un giudizio in base alla media dei voti assegnati), al tracciato di tutte le operazioni svolte su di essi (lettura, commento, valutazione, etc.) indicate con apposite icone. Sulla base del contenuto dei *folder* veniva poi identificato un profilo di utente per raccomandare documenti, persone o comunità pertinenti e avvisi su nuovi oggetti di interesse inseriti nel database.

Tale programma, a cui tuttavia non ha fatto seguito alcuno sviluppo, ha avuto il merito di delineare possibili servizi per potenziare l'utilità degli archivi aperti interpretando i comportamenti di ricerca degli studiosi.

Il meccanismo di gran lunga più importante e ricco di implicazioni per la scoperta di informazione, è tuttavia la possibilità di valorizzare i *link* delle citazioni dei documenti. Da tempo i *citation index* sono riconosciuti come potenti strumenti di ricerca dell'informazione e la trasposizione di questa metodologia sul supporto digitale non

¹³ <<http://oaister.umdl.umich.edu/o/oaister/>>.

¹⁴ <<http://www.openarchives.it/pleiadi/>>

¹⁵ <<http://www.ercim.org/cyclades/>>.

poteva che esaltarne la potenza. Da un punto di vista della struttura dell'informazione, le citazioni rappresentano metadati *impliciti* che permettono nel contesto delle citazioni di una comunità scientifica di scoprire articoli rilevanti su un determinato argomento, di evidenziare l'impatto scientifico di articoli e autori, di ripercorrere l'iter conoscitivo di una scoperta, di valutarne il valore nel tempo. Lo strumento che su web meglio rappresenta queste importanti funzioni è sicuramente "Citebase.IST Scientific Literature Digital Library", sviluppato da Steve Lawrence, Lee Giles, and Kurt Bollacker al NEC Research Institute¹⁶. Esso indicizza 731.912 documenti di cui indica le citazioni, rivela le relazioni con altri articoli (in base a frasi comuni, ad analisi del testo e a co-citazioni), e ne indica l'impatto oltre a permette di dar conto di correzioni o smentite, di impedire la ripetizione di ricerche già fatte, delineare trend di ricerca, aiutare a identificare aree scientifiche emergenti. Il grafo delle citazioni costruito in base al numero di citazioni nel tempo (per anno) delinea a colpo d'occhio la parabola della incidenza culturale di un articolo. Offre inoltre la possibilità di esprimere un giudizio secondo una scala da 1 a 5 e di commentare l'articolo.

Questo modello è stato in parte ripreso dal progetto "Citebase", (copyright di Tim Brody, University of Southampton)¹⁷ ancor oggi (ottobre 2005) in fase di sperimentazione. Il suo campo di applicazione sono al momento alcuni archivi OAI, tra cui "arXiv"¹⁸, e quindi le sue funzionalità non sono ancora completamente apprezzabili, data la assenza di una massa critica di articoli. Oltre a contare le citazioni ricevute fornisce anche il numero di accessi (*hits*) ricevuti da ogni articolo, rappresentati dai clic del mouse per visualizzarli scaricando gli articoli nella memoria *cache* del proprio computer. Si tratta di una grandezza non del tutto obiettiva, data la facilità di manipolazione che rende possibile, ma viene comunque trattata al pari delle citazioni per costruire grafi in relazione col tempo (citazioni e *hits* ricevute nelle singole unità temporali). Oltre a correlare gli articoli tramite le co-citazioni calcola anche una "valutazione di autorevolezza" secondo una metrica sperimentale, non specificata.

SUBJECT GATEWAYS

L'altra tipologia di collezione di risorse che si è ormai affermata nelle biblioteche, sono i cataloghi di risorse di internet. A questo proposito vorrei solo citare un programma per la catalogazione delle risorse di internet (ma si tratta di un sistema abbastanza flessibile per catalogare molte tipologie di materiali: basti pensare che di *default* lo schema di metadati utilizzato è il Dublin Core) che presenta funzionalità interessanti dal punto di vista della costruzione di comunità collaborative di utenti. Il programma si chiama "Scout Portal Toolkit" (SPT)¹⁹ e fa parte di un progetto della University of Wisconsin-

¹⁶ <<http://citebase.ist.psu.edu/>>.

¹⁷ <<http://www.citebase.org/cgi-bin/search>>.

¹⁸ <<http://arxiv.org/>>.

¹⁹ <<http://scout.wisc.edu/Projects/SPT/>>.

Madison e di altre istituzioni culturali americane, per sviluppare strumenti e servizi per trovare, selezionare e presentare in linea informazioni e metadati. Il programma è *open source* e liberamente scaricabile ed è inoltre assistito da un team di esperti.

Il programma è facilmente personalizzabile, sia come apparenza (si può inserire un logo personalizzato, aggiungere news e informazioni, etc. il tutto senza conoscere il linguaggio php) che come struttura dei dati. È OAI compatibile, nel senso che espone i propri metadati agli *harvester* secondo il protocollo OAI-PMH previa registrazione presso il sito dello “Scout portal project”. Utilizza inoltre la tecnologia RSS per distribuire *feed* con le nuove accessioni. Il record è corredato, oltre che dei metadati descrittivi, amministrativi e semantici, anche di uno *screenshot* della risorsa catalogata che, nel caso delle risorse di internet, può essere utilizzato, essendo il *link* normalmente ad una *home page*, per esemplificare un aspetto particolare del servizio (es. una mappa del sito, un menù di opzioni, o la struttura di un record..) o evidenziare un documento pubblicato sul sito.

Tra le funzionalità innovative permette all'utente registrato di assegnare alla risorsa un giudizio lungo una scala da 1 a 5. Questi giudizi vengono utilizzati da un *user agent* implementato nel sistema per “raccomandare” risorse utili che sono naturalmente tanto più efficaci quanto più giudizi l'utente ha espresso. Inoltre utenti autorizzati dall'amministratore del sistema possono inserire commenti, visibili a tutti, alle risorse catalogate. Il sistema dispone anche della possibilità di memorizzare una *query* e di farla rilanciare periodicamente. Il supporto collaborativo è infine sottolineato dalla possibilità di aprire forum di discussione a piacimento.

KM2

Il secondo approccio riguarda l'ingegnerizzazione dei processi, cioè l'estrazione di valore che si ottiene quando processi esistenti sono “anatomizzati” e “ricompilati”. Nel contesto e nel dibattito sulle biblioteche il focus di questo approccio consiste però non tanto nella determinazione di un qualche software migliore, bensì nel tentativo di ridefinire i servizi offerti dalla biblioteca facendo interagire le ormai numerose, ma indipendenti risorse informative cui la biblioteca dà accesso: l'obiettivo è quello di collegare in modo diretto funzioni di ricerca - in cataloghi, banche dati, bibliografie ... senza passare manualmente da una risorsa all'altra -, localizzazione contestuale - cioè relativa alle risorse cui la biblioteca ha accesso - e utilizzo.

Il problema nasce dal fatto che le diverse risorse elettroniche possedute dalle biblioteche sono costruite in primo luogo secondo le esigenze degli editori o aggregatori, senza alcun coordinamento, e anzi in aperta concorrenza, tra di loro. Le diverse risorse adottano software gestionali proprietari non “aperti” e dialoganti, e anzi segreti; i formati utilizzati non sono standard; l'accesso è controllato per ognuna in modo diverso, secondo le scelte del *provider*; come contenuto fanno riferimento ad aggregazioni particolari (es. diverse aggregazioni di riviste) in parte sovrapponibili da editore a editore, spesso ridondanti rispetto alle esigenze delle biblioteche; inoltre sono spesso

poco documentate (nel caso di banche dati di articoli, p. e.: quali riviste vengono spogliate? vengono indicizzati tutti gli articoli o solo i principali? Nel caso di banche dati di riviste elettroniche, con quale criterio è determinata la composizione dei pacchetti di riviste? etc.).

Data l'importanza assunta dall'informazione di internet, poi, va sottolineato che oltre che all'interno della biblioteca (o del consorzio) deve essere assicurata anche l'interoperabilità con gli strumenti di ricerca di rete. Viceversa è già possibile configurare, ad es., Google Scholar e altri servizi di rete, ad es. Connotea (v. avanti), per dialogare con il server che ospita il *link resolver* di un servizio bibliotecario locale: interno ed esterno all'organizzazione, così, si compenetrano.

Si è affermata in tal modo la consapevolezza della necessità di una evoluzione dell'architettura dei servizi elettronici offerti dalla biblioteca, caratterizzata prima come passaggio da ILS (Information Library Services) al LAF (Library Automation Framework)²⁰, o più recentemente come *recombinant library* (Dempsey)²¹.

Da tempo le biblioteche stanno cercando di definire strategie adeguate per aggregare in modo efficace le diverse risorse. Si è cercato di utilizzare il portale della biblioteca per raccogliere in un unico punto l'accesso alle diverse risorse possedute; con software come Z39.50 e sue evoluzioni, si può inviare in sequenza una *query* a diverse banche dati conformandosi ai vari formati e ai diversi meccanismi di ricerca utilizzati; una volta trovata l'informazione dispersa in varie banche dati, la funzione della localizzazione contestuale è affidata alla tecnologia OpenURL grazie alla quale la richiesta di un testo viene gestita da un *link resolver* per identificare se la risorsa cercata è posseduta dalla biblioteca e può essere fornita direttamente all'utente. L'altra tecnologia emergente si basa su OAI PMH per raccogliere i metadati "esposti", cioè resi visibili in rete ai software di indicizzazione, da diversi *data provider* in un deposito unico su cui condurre ricerche avanzate per reperire poi, grazie al solito OpenURL, l'indirizzo alla risorsa pertinente.

Un problema analogo, quello di depositi di informazione non interoperanti e che devono quindi essere utilizzati in modo separato, si evidenzia nel cosiddetto *invisible web*. Come è noto esso riguarda sia documenti in formato non html, sia soprattutto informazione strutturata contenuta in banche dati raggiungibili in Internet, ma ricercabili solo abbandonando l'ambiente di rete e comunicando tramite appositi *gateways* con i software specifici implementati. Tutta questa informazione è irreperibile perché "invisibile" ai motori di ricerca, ma contiene informazioni di grande valore essendo composta in gran parte da banche dati di università, biblioteche, associazioni, agenzie governative, enti pubblici, etc. quindi di materiali strutturati di alto valore informati-

²⁰ A. Rhino, *The end of ILS*, "Inside OLITA", 5, 2001 (Fall/Winter); <<http://www.hpl.hamilton.on.ca/OLITA//InsideOLITA/IO2001No5.htm>>.

²¹ Lorcan Dempsey, *The Recombinant Library: Portals and People*. Pubblicato simultaneamente in "Journal of Library Administration", 39, 4, 2003 e in Sul H. Lee (a cura di), *Improved Access to Information: Portals, Content Selection, and Digital Information*, Binghamton, NY, Haworth, pp. 103-136; <http://www.oclc.org/research/staff/dempsey/recombinant_library/dempsey_recombinant_library.htm>.

vo. Per di più l'accumulo di dati in depositi elettronici è un processo in rapida crescita, assai più rapida di quella delle pagine web.

Nella sua essenza il problema è anche qui quello di rendere possibile la comunicazione tra macchine su di un'unica piattaforma internet, un problema cioè essenzialmente di interoperabilità e quindi di standard, a cominciare dai cosiddetti "servizi di rete", cioè quelle tecnologie come in primo luogo i metadati e lo stesso xml che permettono di portare sulla piattaforma tcp/ip i contenuti delle banche dati e di manipolarli con gli strumenti tipici dei motori di ricerca. Passi avanti sono stati compiuti con l'indicizzazione dei pdf e dei file di microsoft office, ma soprattutto con la conversione dei record delle banche dati in metadati espressi in xml che ne permette l'*harvesting* e l'indicizzazione da parte dei motori di ricerca (OCLC è stato un innovatore in questo campo, facendo indicizzare da Google un subset di record di Worldcat: il servizio si chiama *Find in a library* e permette anche di collegare dati geografici sulla dislocazione delle biblioteche che posseggono il record).

UN NUOVO *BENCHMARK*

Dal punto di vista delle biblioteche e della gestione della conoscenza, però, l'attenzione non va posta sugli aspetti tecnologici (compito di professionalità diverse da quella bibliotecaria), ma soprattutto sui nuovi modelli organizzativi e di servizio agli utenti che tale evoluzione può sviluppare. È quindi di cruciale importanza sottolineare che l'evoluzione dei servizi in internet tende a andare anche al di là dell'obiettivo sopra ricordato di riconnettere in un continuum le funzioni di ricerca, localizzazione e utilizzo. Si deve anzi proprio a servizi commerciali in internet come Amazon, Google, Yahoo ed eBay, per citare i più importanti, la creazione di una serie di risposte alle esigenze degli utenti che prefigurano un vero e proprio *benchmark* per la costruzione dei servizi bibliografici.

Le tappe di questa evoluzione sono state ripercorse nel *blog* di Lorcan Dempsey²², vicepresidente di OCLC, sotto la parola chiave di *Amazoogles* e recentemente illustrate anche da Roy Tennant che in un intervento al convegno di LITA del settembre 2005, ripercorre le tappe dell'evoluzione del web negli ultimi dieci anni da Altavista a quello che, con analogia combinazione dei nomi di Google e Amazon, chiama, *Googlezon*²³. Il problema è presentato in forma schematica e chiara dalle White papers del progetto "Silkworm" (baco da seta) del provider di servizi per biblioteche Talis – autore anche dell'innovativo software "RedLightGreen" per far comunicare le risorse delle biblioteche aderenti a RLG – per la definizione di specifiche e modelli per servizi di automazione delle biblioteche basati sull'interoperabilità in un contesto di rete. Ecco una sintesi dei risultati del confronto basato soprattutto tra Amazon e l'OPAC:

²² <<http://orweblog.oclc.org/>>.

²³ Roy Tennat, *Googlezon, Episode VI: Return of the Librarians*, in LITA National Forum 2005, San Jose, CA; <<http://www.ala.org/ala/lita/litaevents/litanationalforum2005sanjoseca/nf2005keynote.htm>>.

Caratteristica	Amazoogole	Opac
Onnicomprensività	Google abbraccia tutto il web; Amazon il mercato dei libri nuovi e di seconda mano, eBay oggetti da vendere all'asta. Non c'è bisogno di andare altrove.	L'utente vede solo ciò che si trova nella biblioteca. Se non trova il libro deve cercarlo da qualche altra parte. Non c'è modo di sapere se un testo non c'è perché la biblioteca non lo possiede o perché non esiste.
Senza pareti	Non ci sono frustranti barriere nel passaggio da un servizio all'altro.	L'uso di diversi servizi richiede continui log out e log in, spesso con diverso account.
Immediatezza	Gli utenti di questi siti sono soddisfatti rapidamente e con facilità, con un click.	La biblioteca non offre spesso alcuna opzione per ottenere subito il libro o l'articolo.
Ricchezza	Vengono offerti numerosi servizi: recensioni, giudizi, visualizzazione della copertina, interviste all'autore, raccomandazioni di testi simili; particolarmente utile la ricerca di parole nel testo dei libri (search inside di Amazon e Google Print).	La biblioteca offre solo informazioni bibliografiche e sulla collocazione.
Partecipazione	L'utente è incoraggiato a scrivere recensioni e guide, a dare il suo feedback e valutazione, a far parte di liste di discussione. eBay è quasi del tutto basata sulla partecipazione.	Non esiste alcun modo di partecipare all'OPAC della biblioteca.
Personalizzazione	Es. Amazon suggerisce libri sulla base di quello che l'utente ha comprato.	Manca del tutto nel catalogo OPAC.
Soddisfazione	I servizi realizzano la soddisfazione dell'utente esperto e inesperto.	Gli OPAC soddisfano al meglio le esigenze di utenti esperti e consapevoli di cosa vogliono trovare (funzione find, distinta da search).

KM3

Il terzo approccio del *Knowledge Management* ai problemi della creazione e distribuzione della conoscenza riguarda l'esternalizzazione e condivisione della conoscenza implicita. Il focus è sulla comunicazione interpersonale, sulla condivisione e sulla collaborazione che si realizza nelle comunità di pratica o epistemiche in cui si attua un

apprendimento che non è frutto di un disegno cosciente o di una razionalità riconoscibile, bensì di strutture emergenti e significati nuovi.

Già nella prima parte di questo scritto si sono portati alcuni esempi esaminando alcuni servizi implementati per OAI dai service provider. Gli strumenti di queste attività sono però in continuo aumento. Dopo la posta elettronica, sempre più assediata dallo *spam* e dalle minacce degli *hacker*, sono venuti in primo piano gli RSS come strumento per comunicare informazione; anche i *blog*, sono strumenti di socializzazione perché sono spesso linkati l'uno con l'altro e ospitano commenti. Tramite i *blog* si ha una rapida condivisione di spunti e indicazioni che rendono possibili inattese connessioni e percorsi originali, rendendo l'insieme dei *blog* (*blogsphere*) un enorme deposito di conoscenza. È questa la riflessione che ha portato David Sifry, a seguito della lettura del "Cluetrain manifesto"²⁴, a costruire un motore di ricerca per i *blog*, con lo scopo di valorizzare l'informazione che emerge dalle conversazioni. Questo motore, che si trova sul sito di "Technorati", applica il criterio della popolarità nel *ranking* dei risultati e si applica alla segnalazione delle variazioni registrate dai *permalink* che si verificano nella *blogsphere*; la ricerca avviene per parole chiave e per *tag* inseriti dai *blogger* stessi. In tal modo Sifry contrappone ad una metafora della rete come "biblioteca" la visione di essa come un grande "fiume di conversazioni" alimentato da innumerevoli rivoli²⁵.

Un servizio da citare nell'ambito degli strumenti collaborativi è "wiki wiki", un editor di rete che rende possibile pubblicare rapidamente testi in Internet. Il suo uso è molto frequente in ambienti collaborativi controllati per condividere, aggiornare o coeditare testi. La sua efficacia si manifesta anche in quel monumento, nel bene e nel male, alla libera collaborazione che è "Wikipedia"²⁶, l'enciclopedia ad accesso gratuito su web.

Questi diversi ambienti di comunicazione, cooperazione, condivisione costituiscono uno dei più interessanti fenomeni del web negli ultimi anni per quanto attiene alla produzione di conoscenza.

FOLKSONOMY

Recentemente si sono diffusi sul modello di "Flickr" e "del.icio.us" – siti per la condivisione di fotografie l'uno e di *bookmark* il secondo –, servizi come "Connotea" del Nature Publishing Group (NPG) o "CiteUlike", pensati per condividere documenti in ambito accademico. Si tratta in pratica di strumenti per organizzare i propri siti preferiti in rete e per dividerli. Ciò che contraddistingue questi servizi è anche il modo in cui le risorse sono organizzate, cioè attraverso parole chiave tratte dal linguaggio naturale dell'utente, dette *tags*. L'uso di *tags*, indispensabile per identificare il contenuto di oggetti non testuali, come le fotografie, è invece meno ovvio per i *bookmark* o la posta

²⁴ Le tesi del "Cluetrain Manifesto" si leggono in italiano all'indirizzo: <<http://www.mestierediscrivere.com/testi/Tesi.htm>>.

²⁵ *Face value. The life and soul of the internet party*, 6 ottobre 2005, from "The Economist" print edition, in: <http://www.economist.com/displaystory.cfm?story_id=4484119>.

²⁶ <<http://www.wikipedia.org/>>.

elettronica. Questo metodo è stato designato col termine *folksonomy*, dalla contrazione delle parole *folks* – gente – e *taxonomy* – tassonomia, da Thomas Vander Wal che descrive questo processo come “classificazione sociale dal basso”²⁷. In realtà non si tratta di una nuova tassonomia (anche se c’è chi suggerisce un uso codificato di semplice punteggiatura per inserire relazioni semantiche elementari: ad es. = per sinonimi) ed è forse preferibile l’espressione “*social bookmarking*” o “*collaborative tagging*” per definire questa pratica. Le sue caratteristiche sono quindi la costruzione dal basso, l’assenza di struttura gerarchica, il loro uso in un contesto sociale.

A mia conoscenza il primo servizio che ha utilizzato i *tag*, chiamandoli etichette (*label*), per organizzare i contenuti, è stata Gmail, il servizio di e-mail di Google che si caratterizza per l’ambizioso progetto di costituire con le mail inviate e ricevute da soggetti selezionati (la casella di Gmail si otteneva solo per invito) una banca dati da cui estrarre con le tecniche di ricerca di Google, la conoscenza esplicita ed implicita che si nasconde nella comunicazione interpersonale: per la prima volta un *provider* invitava a non cancellare e anzi ad archiviare le mail e offriva a tale scopo gratuitamente lo spazio enorme di un gigabyte di ampiezza (oggi sono quasi 3). Le mail erano poi organizzate non in cartelle e sottocartelle bensì appunto etichette (*label*) e raggruppate insieme alle rispettive risposte per essere visualizzate “sotto forma di conversazione”.

L’interesse di questo approccio nasce però dal fatto che ancora una volta il tentativo di dominare l’informazione di rete punta a modificare il rapporto tra i tradizionali protagonisti dell’informazione: autore, lettore e mediatore. All’inizio la comunità che aveva definito il Dublin Core si era rivolta direttamente agli autori dei documenti web perché inserissero nei loro file dei metadati, cioè una struttura informativa più semplice e agile del record catalografico basato sul MARC, che potesse poi essere utilizzata a fini di ricerca dai motori di internet. Con l’utilizzo di *tag* degli utenti per classificare i contenuti – siano essi fotografie, siti web, libri o articoli di riviste – la mediazione di terzi tra autore e lettore, subisce un ulteriore colpo e viene in primo piano il lettore, cioè il destinatario dell’informazione.

La struttura del web basata sulla presentazione tipografica dei documenti aveva inizialmente riproposto il modello bibliotecario di organizzazione dell’informazione tanto più che si riteneva prioritaria la selezione delle risorse di rete in base a criteri di qualità posti, per così dire, dall’esterno e provenienti dal mondo delle pubblicazioni tradizionali²⁸. L’inadeguatezza degli strumenti tradizionali dell’organizzazione bibliotecaria, troppo complessi e costosi per svolgere il loro compito nel contesto della rete, ma anche – è la tesi che qui si cerca di motivare – spesso inadeguati a cogliere l’essenza delle nuove forme di conoscenza che si costruiscono e veicolano in rete, spingono

²⁷ Tony Hammond, Timo Hannay, Ben Lund, and Joanna Scott, *Social Bookmarking Tools (I)*, *A General Review*, in “D-Lib Magazine”, 11, 4, 2005; <<http://www.dlib.org/dlib/april05/hammond/04hammond.html>>.

²⁸ Per una discussione del concetto di autorità, in riferimento ai sostenitori del *social software* si veda la seguente colonna di Peter Morville, *Authority*, October 11, 2005, in “Semantic studios”; <<http://semanticstudios.com/publications/semantics/000057.php>>.

alla ricerca di nuove soluzioni in parte, come si è visto, del tutto esterne ai tradizionali meccanismi di organizzazione dell'informazione e della conoscenza, in parte che ad essi necessariamente ritornano. La risultante di questa dislocazione di approccio è una nuova centralità dei valori e dei significati, cioè una nuova idea di web semantico.

La *folksonomy* rappresenta in questa evoluzione una spinta verso l'organizzazione dal basso (*bottom up*) delle risorse di rete, creata socialmente e tipicamente piatta, cioè non gerarchica. Come è stato sottolineato questa soluzione nasce già viziata da incongruenze e difetti di fondo, prima di tutto la mancanza di universalità, ovvero da un marcato soggettivismo esemplificato dall'uso di *tag* come “*me*” in Flickr o “*toread*” (“daleggere”) per non parlare della mancanza di gestione di sinonimi, di parole polisemiche, etc. D'altro lato sono noti anche i difetti delle classificazioni e ontologie costruite a priori secondo criteri razionali: non reggono alla prova del tempo e necessitano un costante e costoso aggiornamento che li allontana dal mondo reale, diventano sempre meno comprensibili per gli utenti, in tal modo si fanno sempre di più “verità rivelata”, e richiedono una qualche “casta sacerdotale” per tramandarsi; non sono scalabili, né dal punto di vista della corrispondenza col mondo reale (cioè più oggetti si devono classificare, più le tassonomie perdono di precisione), né da quello della coerenza interna (ad esempio la struttura gerarchia non riesce a rendere conto dei diversi approcci disciplinari, da cui ad esempio il ricorso a “faccette” per attenuare la struttura gerarchica); infine, ma il discorso sarebbe assai complesso, non sono condivise da tutti, cioè non tengono conto delle diversità di senso, delle preferenze, degli obiettivi etc. dei diversi soggetti, in una parola dei significati che ognuno di noi inevitabilmente attribuisce ai contenuti²⁹.

Ma ciò che cambia le carte in tavola è il fatto che la rete non è solo una piattaforma di scambio neutrale e piatta ma si articola sempre più in comunità, cioè in aggregazioni spontanee caratterizzate da comuni interessi, da valori condivisi e da un comune linguaggio: sono queste comunità che attribuiscono *senso* agli eventi della rete e articolano relazioni semantiche. In altri termini la rete è luogo di una continua negoziazione e definizione delle condizioni e delle modalità dello scambio di informazione e di conoscenza: i *tag* attribuiti ad un articolo sono confrontati con quelli dati da altri utenti, fino a far prevalere quello che soddisfa le esigenze della maggior parte dei partecipanti e che quindi viene di fatto accettato come più appropriato.

Questa negoziazione semantica è essenziale in un ambito di condivisione e collaborazione, e questa è proprio una delle prime funzioni che conferiscono alla comunità la capacità di socializzare per produrre conoscenza e apprendere³⁰. Lungo questo ra-

²⁹ Alcuni di questi aspetti sono discussi in: Scott A. Golder and Bernardo A. Huberman, *The Structure of Collaborative Tagging Systems*, in <<http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf>>. Lo studio proviene da HP labs. Tim O'Reilly nel suo confronto tra web 1.0 e web 2.0 sembra predire la fine delle tassonomie in favore delle *folksonomies*; v. *What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software*, by Tim O'Reilly (09/30/2005); <<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>>.

³⁰ Bernard Ancori, Antoine Bureth and Patrick Cohendet, *The economics of knowledge: the debate about codification and tacit knowledge*, in “Industrial and corporate change”, 9, 2, 2000, pp. 255-287.

gionamento vi è chi suggerisce che il *social bookmarking* possa costituire un modo per diminuire la distanza tra sistemi di classificazione consolidati e utenti per costituire un accesso in linguaggio naturale alle classificazioni formali o per indicare *wished path*, percorsi semantici desiderati, che modificando il disegno originale divenuto obsoleto, suggeriscano termini e approcci considerati dagli utenti più idonei per identificare un determinato argomento³¹. Alcuni studi suggeriscono poi che i *tags* degli utenti siano particolarmente utili in campi della conoscenza non ancora maturi in cui devono ancora emergere e consolidarsi modelli di relazioni stabili tra le varie entità³².

CONCLUSIONI

L'approccio del KM ai problemi di organizzazione della conoscenza ci ha permesso di evidenziare e articolare in modo funzionale una serie di fenomeni e di nuovi servizi che si vanno sviluppando nel mondo dell'informazione integrando e trasformando i tradizionali strumenti bibliotecari per l'accesso alla conoscenza. La questione è se sia compito delle biblioteche occuparsi anche di questo tipo di informazione o se non si debba dare per scontata una divaricazione nel mondo dell'informazione che sembra nascere da una frattura epistemologica che proprio il KM aiuta a identificare. Si tratta di una posizione rassicurante ma anche pericolosa perché potrebbe avere l'effetto di condannare le biblioteche alla marginalità o a ridurle a compiti di conservazione.

È invece assai più probabile, sul breve periodo, che i servizi organizzati da operatori commerciali che agiscono in rete, influenzeranno sempre di più il modo di "fare informazione" delle biblioteche che, come ogni altra organizzazione che opera nel mercato globale, tenderà a evolversi verso la fisionomia competitiva della *learning organization*³³ presentandosi quindi sempre meno come insieme di tecniche e strumenti, e anzi, per dirla con Enzo Rullani, sciogliendosi "in processi che sono retti dalle competenze delle persone e dalle *capabilities* dinamiche di gruppi di persone (...). Il tratto distintivo dei diversi concorrenti sul mercato è dato non più dallo stabilimento, dal prodotto o dal nome, come era una volta, ma dalle *core competencies*, dalle competenze distintive cui ha accesso. Non conta cosa possiede ma che cosa sa fare". In tal modo i "flussi comunicativi in entrata e in uscita trasformano la trama rigida delle strutture organizzative, tessendo e ritessendo continuamente il disegno e il pensiero della *corporation*. La tecnologia diventa *know-logy* (Dricker, 1968) sommando gli aspetti formali del sapere astratto con quelli informali, pratici, del saper fare"³⁴.

³¹ Peter Merholz, *Metadati for the masses*, Adaptive path, October 19, 2004; <<http://www.adaptivepath.com/publications/essays/archives/000361.php>>.

³² David N. Sturtz, *Communal categorization: the folksonomy*; <<http://www.davidsturtz.com/drexel/622/communal-categorization-the-folksonomy.html>>.

³³ Peter Senge, *The fifth discipline*, 1990.

³⁴ Enzo Rullani, *Economia della conoscenza. Creatività e valore nel capitalismo delle reti*, Roma, Carocci, 2004, p. 47.