

Connettere la conoscenza: quattro livelli di interoperabilità*

CLAUDIO GNOLI

Biblioteca della scienza e della tecnica
Università di Pavia
claudio.gnoli@unipv.it

La connessione delle informazioni in rete è senza dubbio uno dei maggiori paradigmi nelle scienze dell'informazione contemporanea. La svolta è cominciata con l'apertura di Internet, nato in ambito militare, alle università americane e ai personal computer; è continuata con l'invenzione, giusto un quarto di secolo fa, delle interfacce ipertestuali e grafiche del World Wide Web; quindi con il coinvolgimento di fasce di popolazione sempre più ampie in reti *sociali* (Web 2.0) di diari, di indirizzi di risorse, di fotografie, di video, di amicizie, di recensioni, di conoscenze enciclopediche e ora anche di pubblicazioni accademiche.

Quella della rete è la nuova grande metafora della conoscenza, che si affianca e sovrappone alle metafore classiche del *cerchio* delle materie di apprendimento, della *scala* di gradini dalla crescente nobiltà ontologica, dell'*albero* in cui tutte le categorie sono ricondotte a una gerarchia di appartenenza, della *mappa* che codifica e rappresenta le relazioni spaziali di un territorio, geografico o concettuale che sia [Barsanti 1992; Santoro 2002; Weingart 2013; Lima 2014]. Rispetto ai precedenti, la rete (ovvero rizoma o, nel nome del protocollo di trasmissione principe, ragnatela) è un tipo di grafo che lascia più libertà di connessione da un qualsiasi nodo potenzialmente a un qualsiasi altro, nell'una o nell'altra direzione. Una tale struttura, come ben sanno gli studiosi del cervello, è in grado di far emergere percorsi e proprietà di una molto maggiore complessità e sofisticazione rispetto ai grafi di tipo più costrittivo (orientati, aciclici ecc.) [Wilson 1978].

Una maggiore complessità che significa anche maggiore caos rispetto all'ideale tradizionale del controllo bibliografico, poiché le connessioni sono in continuo sviluppo e spesso transitorie; svantaggi indubbi, che nel complesso però sono forse superati dai vantaggi in termini di potenziale informativo: per Weinberger [2010] non dovremmo allarmarci troppo dell'avanzare di questo "disordine", dal momento che le interfacce digitali e i motori di ricerca renderebbero meno indispensabili gli sforzi tradizionali di catalogazione, con buona pace dei bibliotecari all'antica e degli appassionati dell'ordinamento sistematico delle conoscenze come Mortimer Adler, che ne propugnò l'applicazione nell'*Enciclopedia britannica*. La successiva tappa (Web 3.0) di questa evoluzione nel collegamento delle informazioni è stata identificata nella loro associazione con metadati e schemi per il riconoscimento automatico di connessioni; essa si basa su un'identificazione univoca (*Uniform resource identifier*, URI) non più solo degli interi documenti, tipicamente in forma di file, ma di ogni dato elementare, come il fatto che Collagna si trovi a 830 metri di altitudine o il fatto che un lupo sia stato avvistato ieri a Collagna: il potersi riferire a essi a partire da qualunque altra risorsa apre infatti la possibilità di stabilire connessioni significative fra gli stessi contenuti, ad esempio deducendo automaticamente che un lupo sia stato avvistato a 830 metri di altitudine e inserendo questa informazione nel nuovo contesto di una statistica delle abitudini dei lupi, formando progressivamente un *web dei dati* o *web semantico*.

Il ruolo ineludibile dell'organizzazione della conoscenza (KO)

La crescente interoperabilità fra i dati è certo un avanzamento molto rilevante, dal punto di vista tecnico, per chi si interessa della diffusione e dell'uso delle in-

* Testo rivisto dell'intervento tenuto in occasione del convegno "La biblioteca connessa. Come cambiano le strategie di servizio al tempo dei social network" (Milano, Palazzo delle Stelline, 13-14 marzo 2014). Gli indirizzi web citati sono stati verificati il 2 giugno 2014.

formazioni. Tuttavia, come sempre, la tecnica non basta. Perché sia giustificato l'aggettivo *semantico* con cui è ormai uso descrivere questo tipo di applicazioni, occorre che i dati siano interoperabili non solo in quanto stringhe di bit ma anche in quanto significati. Occorre cioè riconoscere e trattare quei classici problemi di omografia e sinonimia fra termini e di relazioni fra concetti che sono noti da molto tempo ai bibliotecari e ai documentalisti, ma solo recentemente stanno giungendo anche all'attenzione degli sviluppatori di risorse sul web.

I bibliotecari hanno tradizionalmente e utilmente adottato una distinzione fra catalogazione descrittiva e catalogazione per soggetto, ossia riferita appunto ai contenuti conoscitivi dei documenti. In Italia abbiamo talvolta chiamato quest'ultima anche catalogazione *semantica*, ad esempio auspicando che siano "semantici" i cataloghi nell'offrire adeguati strumenti di ricerca e presentazione delle registrazioni bibliografiche (in base alle stringhe di soggetto e alle classi attribuite ai documenti) [Gnoli et al. 2004]. Nella letteratura anglofona tuttavia questo termine non è utilizzato, e "indicizzazione semantica" va tradotto esclusivamente con *subject indexing* per non venire fraintesi dai colleghi stranieri. A maggior ragione ora che a *semantic* viene associato il senso di dati di qualunque natura registrati nei formati di scambio tipici del Web 3.0, mi convinco che ci convenga abbandonare del tutto il vecchio utilizzo di *indicizzazione semantica* e sforzarci di adottare solo il classico e univoco *indicizzazione per soggetto*.

In un ambito più generale di quello bibliotecario, e comprendente lo studio e il trattamento delle conoscenze anche in senso storico, sociale o informatico, l'identificazione e la connessione dei contenuti conoscitivi sono state espresse negli ultimi decenni con il termine *organizzazione della conoscenza*. Già utilizzato negli importanti testi di H.E. Bliss degli anni Trenta, questo termine è stato poi adottato sistematicamente a partire dagli anni Settanta con gli studi specifici della tedesca Ingetraut Dahlberg, il cui *Wissensordnung* è stato tradotto con *knowledge organization* (KO) e in italiano appunto con *organizzazione della conoscenza* – o anche *delle conoscenze* o, nell'elegante soluzione di Giliola Negrini, prima seguace italiana di Dahlberg, *del sapere*.

Ci teniamo a parlare di "conoscenze" e non solo di "informazioni" in quanto le pensiamo come sistemi di informazioni singole ampi e coerenti, come quelli adottati nell'organizzare e sviluppare intere biblioteche, la cui caratteristica sia quella di essere – appunto – con-

nesse. Peraltro l'espressione *information organization*, adottata in alcuni dipartimenti universitari americani, è intesa in un senso sostanzialmente equivalente a *knowledge organization*, come ha mostrato Hjørland [2012].

Il ricorso a queste nozioni, d'altronde, è concettualmente inevitabile: nel momento in cui si voglia connettere dei contenuti conoscitivi in strutture coerenti, si sta facendo per definizione organizzazione della conoscenza, come la facevano già Lullo, Bacone, Linné e Otlet. Purtroppo non sempre si è consapevoli, ed è per questo che è utile e necessario identificare principi e tecniche in una branca specifica che adotti una propria terminologia coerente: solo in questo modo si possono contenere gli sprechi delle scoperte dell'acqua calda, tanto frequenti in questo campo spesso venduto per nuovo con nomi trendy, o all'inverso dato troppo per scontato, forse proprio a causa della sua natura fondamentale e intrinseca alle strutture stesse del pensiero umano.

Ai sistemi per l'organizzazione di conoscenze, sempre più spesso sintetizzati nella pratica sigla KOS (Knowledge organization system), appartengono tradizionali strumenti bibliotecari come i soggetti, i tesauri e gli schemi di classificazione; ma anche le gerarchie degli oggetti di studio delle scienze naturali o delle voci nei menù di navigazione di un sito, spesso chiamate tassonomie [Hedden 2014]; i sistemi di tag attribuiti dagli utenti del Web 2.0 o folksonomie; le mappe tematiche (*topic maps*) adottate per la visualizzazione grafica di connessioni fra termini; le terminologie studiate dai lessicografi, dagli esperti di trattamento automatico delle lingue e da quelli di singoli domini specialistici di conoscenza; le ontologie teorizzate dai logici con propensione per le applicazioni nelle scienze dell'informazione e sviluppate da chi si occupa di intelligenza artificiale e ora anche di web [Biagetti 2010], dei quali l'Italia vanta a Padova e Trento importanti scuole di livello internazionale.

Raggruppare questi strumenti eterogenei nella nozione unificante di organizzazione della conoscenza non significa categorizzarli secondo un interesse personale, che può intersecarsi con altri interessi e categorie: è invece un'opportunità data dall'effettiva comunanza dei principi fondamentali con cui i concetti possono venire espressi e collegati, mediante forme e relazioni standardizzate come quelle impiegate nei tesauri o nell'analisi a faccette, che risultano efficaci per una grande varietà di applicazioni. Già nel suo eccellente manuale Foskett

[2001] fa osservare come, al di là dei gerghi particolari e della diversità di presentazione degli indici, i principi alla base dei sistemi verbali di indicizzazione per soggetto e degli schemi di classificazione siano sostanzialmente gli stessi: e ciò vale anche per gli altri tipi di KOS che abbiamo rapidamente ricordato.

In questo contesto allargato, la distinzione fra indicizzazione descrittiva e per soggetto tende a diventare sempre meno centrale, tanto che molti informatici faticano ad afferrarne il senso: nel web dei dati titoli, autori ed editori sono infatti semplicemente attributi (i *library linked data*, LLD) di particolari classi di oggetti, quali i libri, che si possono descrivere con ontologie apposite al pari di qualunque altra classe di oggetti [Giunchiglia et al. 2014]. Anche chi si occupa dei modelli concettuali delle scienze dell'informazione comincia a pensare che i metadati descrittivi si possano trattare come faccette dedicate alla forma e al supporto nell'ambito di KOS più generali [Smiraglia 2012]. L'organizzazione della conoscenza potrebbe cioè esprimere contemporaneamente diverse dimensioni: i fenomeni oggetto di studio, le prospettive con cui vengono trattati e appunto i supporti che ne veicolano la trattazione; un KOS ideale dovrebbe allora permettere di cercare separatamente la trattazione di un certo fenomeno in qualsiasi prospettiva e supporto, oppure una certa prospettiva applicata a qualsiasi fenomeno ecc.

Connettere concetti

I KOS dunque servono a rappresentare le connessioni dei concetti, possibilmente espressi da termini controllati o da notazioni di classe. Naturalmente, come non mancano di sottolineare fino alla noia alcuni studiosi di organizzazione della conoscenza, né le connessioni né i termini adottati da un KOS sono mai del tutto neutrali, ma dipendono più o meno consapevolmente dalle particolari prospettive dei suoi creatori e utilizzatori. Un ottimo esempio di questo problema rimane l'analisi del termine *sviluppo* nei tesauri delle organizzazioni internazionali, che ha evidenziato come essi lo sussumano regolarmente alle categorie dell'economia, trascurandone invece i sensi sociali e umani e tradendo così l'implicita adozione di una prospettiva capitalista [Severino 2006].

La reazione a questa presa di coscienza può incanalarsi lungo due strade. Da un lato, il riconoscimento di un numero maggiore di prospettive e di possibili sche-

mi mentali rispetto a un unico "lato occidentale delle cose" [Fossati 1996] è un valore per il mondo plurale della contemporaneità. È giusto cioè affiancare alla garanzia bibliografica (*literary warrant*) una garanzia culturale (*cultural warrant*), che assicuri il rispetto delle prospettive di comunità diverse per estrazione sociale, ubicazione geografica e tradizioni [Beghtol 1986]. L'approccio dell'analisi di dominio, sostenuto specialmente da Birger Hjørland, teorizza che i diversi approcci con cui comunità di ricerca diverse (diciamo i fitogeografi e gli etnobotanici) organizzano le conoscenze degli stessi oggetti vadano esplicitamente riflessi in KOS diversi, ciascuno espressione di una determinata prospettiva.

Questa scelta, se da un lato garantisce una posizione politicamente corretta a chi la teorizza, nella pratica però entra in conflitto con le esigenze dell'interoperabilità: i KOS diversi utilizzati da ciascuna comunità infatti, seppur internamente coerenti, saranno verso l'esterno tutt'altro che connessi, proprio per la loro dichiarata incommensurabilità culturale.

Si tratta del resto di un problema di lunga data in epistemologia e filosofia del linguaggio: i diversi paradigmi che coesistono in culture diverse o che si succedono al momento delle rivoluzioni scientifiche sono completamente sconnessi l'uno dall'altro? Qualsiasi "traduttore" di una terminologia in un'altra è in realtà un "traditore"? Il fatto che culture umane anche molto lontane, come quella occidentale e quella cinese lungo la medievale Via della seta, stimolate dall'opportunità del commercio, abbiano comunque trovato delle basi comuni per intendersi in qualche modo suggerisce che le barriere non siano poi così assolute. È vero che il campo semantico di un termine spesso non è esattamente sovrapposto a quello dei suoi presunti equivalenti in altre lingue (*wood* in inglese significa contemporaneamente "bosco" e "legno", mentre esprimendosi in italiano occorre scegliere l'uno o l'altro); ma nella pratica la possibilità delle perifrasi, e nelle ontologie quella delle definizioni esplicite, permettono di aggirare gli ostacoli.

Muovendosi dalla nobiltà delle teorie alla necessità della pratica, l'insegnamento che queste ricerche ci trasmettono è che per evitare ambiguità i concetti dovrebbero essere il più possibile definiti esplicitamente, appunto per mezzo di connessioni con altri concetti codificati nelle relazioni standard dei KOS; e che a sua volta la prospettiva adottata da ogni KOS dovrebbe essere esplicitamente dichiarata.

Connettere discipline

Uno dei modi in cui le comunità di ricerca esprimono il proprio approccio distintivo è la suddivisione della conoscenza per discipline. KOS come le classificazioni bibliografiche o i settori scientifico-disciplinari in base ai quali il Ministero dell'istruzione articola la valutazione di progetti e documenti sono impostati su classi di natura disciplinare. Per indicizzare i contenuti di un documento occorre quindi decidere innanzitutto se rientrano nell'una o nell'altra disciplina: se per esempio si tratti di un testo di diritto dell'ambiente oppure di ecologia oppure di architettura del paesaggio, alternative che nella Classificazione decimale Dewey corrispondono a tre punti assai lontani fra loro.

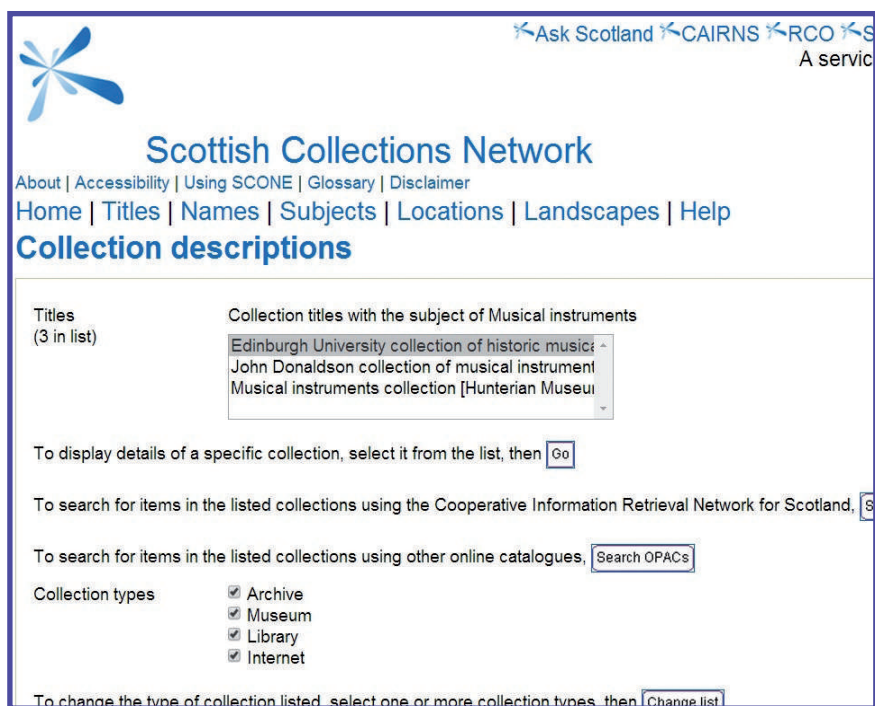
L'assunto soggiacente a questa impostazione è che quando uno studioso produce un documento e quando un lettore lo utilizza lo facciano sempre da una ben determinata prospettiva disciplinare, sicché risulterà loro più utile connettere il documento ad altri che applichino la stessa prospettiva a qualsivoglia oggetto di studio che connetterlo invece a documenti che condividano l'oggetto di studio (*fenomeno*) ma lo trattino con approcci diversi. Tale filosofia dell'indicizzazione è quella tutt'ora insegnata nei manuali di classificazione bibliografica, e si fonda certo su esperienze compiute nei servizi bibliografici degli scorsi decenni. Essa assegna all'organizzazione della conoscenza un ruolo di conservazione dello status quo disciplinare: la funzione dei cataloghi sarebbe quella di rappresentare il modo in cui gli studiosi hanno suddiviso il sapere nel passato e fino a ora, e il ruolo del lettore quello di apprendere tali suddivisioni e conformarvisi rispettosamente.

Alcuni tuttavia suggeriscono che l'organizzazione della conoscenza possa assumere anche un ruolo meno passivo e più efficace nello stimolare nuove ricerche, nuove connessioni fra i contenuti del sapere reso disponibile da biblioteche, archivi e musei. Non è detto infatti che gli autori, nel fornire e presentare le loro conoscenze su qualcosa, ne abbiano esaurito tutte le possibili interpretazioni e connessioni. Spesso anzi accade che fra

documenti di discipline diverse esistano connessioni potenziali che non sono ancora state notate, ma che per il solo fatto di essere identificate costituirebbero un avanzamento della conoscenza – una nuova fruttuosa connessione nel grande grafo del sapere umano [Davies 1990; Beghtol 1995; Revelli 2004].

Rick Szostak, docente di economia alla University of Alberta rivoltosi anche all'organizzazione della conoscenza, osserva come questa identificazione di nuove connessioni fra fenomeni già noti sia un modo di avanzamento del sapere più frequente che la descrizione di fenomeni autenticamente nuovi: ormai sappiamo già moltissime cose sui fenomeni del mondo (è molto raro per esempio che scopriamo una nuova specie di mammiferi), mentre lo spazio nel quale abbiamo ampi margini di progresso è la loro spiegazione con teorie e metodi diversi da quelli con cui li abbiamo considerati finora, che stabiliscano nuove connessioni causali con altri fenomeni (quali fattori antropici influenzano la sopravvivenza del nostro mammifero? Oppure, quale influenza hanno l'ambiente familiare e la discendenza genetica sul tasso con cui in età adulta si manifesta una certa patologia?) [Szostak 2008].

I KOS basati sulle discipline confinano inevitabilmente queste connessioni entro ambiti ristretti: è possibile infatti stabilirne soltanto all'interno di ciascuna di-



The screenshot shows the 'Scottish Collections Network' website. At the top right, there are logos for 'Ask Scotland', 'CAIRNS', 'RCO', and 'A service'. The main heading is 'Scottish Collections Network'. Below it, there are navigation links: 'About | Accessibility | Using SCONE | Glossary | Disclaimer', 'Home | Titles | Names | Subjects | Locations | Landscapes | Help', and 'Collection descriptions'. The main content area is titled 'Titles (3 in list)' and shows a list of collection titles with the subject of 'Musical instruments'. The list includes: 'Edinburgh University collection of historic music', 'John Donaldson collection of musical instrument', and 'Musical instruments collection [Hunterian Museum]'. Below the list, there are instructions: 'To display details of a specific collection, select it from the list, then [Go]', 'To search for items in the listed collections using the Cooperative Information Retrieval Network for Scotland, [S]', and 'To search for items in the listed collections using other online catalogues, [Search OPACs]'. There is also a section for 'Collection types' with checkboxes for 'Archive', 'Museum', 'Library', and 'Internet'. At the bottom, there is a link: 'To change the type of collection listed, select one or more collection types, then [Change list]'.

Ricerca contemporanea in archivi, biblioteche, musei e Internet di collezioni riguardanti gli strumenti musicali nello Scottish Collections Network

sciplina, ma non tra discipline diverse (faccette ma non relazioni di fase, nella terminologia di Ranganathan). Classificazioni come la Dewey e la Bliss offrono poche o nulle possibilità di esprimere relazioni fra discipline diverse, istruendo anzi esplicitamente i classificatori a scegliere la disciplina “prevalente” nel documento e a trascurare le altre. Altri sistemi come la CDU e la Colon offrono possibilità combinatorie assai maggiori (a ciò si riferisce l’immagine del Meccano che ispirò a Ranganathan l’idea delle faccette); ma il significato delle componenti che esse possono connettere rimane comunque vincolato alla rispettiva definizione all’interno di una disciplina: è possibile esprimere cose come “diritto forestale in relazione alle scienze politiche”, ma non tout court “foreste in relazione agli enti locali”, perché le classi sono definite come suddivisioni di discipline e non di fenomeni. La libertà di connessione è quindi solo parziale: è impossibile effettuare una ricerca bibliografica che produca un insieme di documenti sulle foreste trattate da qualsiasi prospettiva, giuridica o amministrativa o urbanistica o biologica. Se apertissimo i KOS anche a queste connessioni, si è sostenuto fino a ora, essi genererebbero troppo rumore, suggerendo al giurista anche informazioni botaniche e gestionali che non gli servono.

Ma è davvero così? Sono sempre più numerosi i convegni, i progetti e le associazioni che dichiarano invece di rivolgersi a studi “interdisciplinari”, a una conoscenza “transdisciplinare”, a studi “integrati”. Anche l’organizzazione della conoscenza dovrebbe offrire loro strumenti flessibili che siano di stimolo anziché di ostacolo allo sviluppo di nuove connessioni. È quanto ha dichiarato esplicitamente il Manifesto di León, sottoscritto da alcuni partecipanti all’ottavo congresso della sezione spagnola dell’International society for knowledge organization (ISKO): i nuovi KOS dovrebbero essere definiti a partire dai fenomeni, anziché dalle discipline, e rendere possibile una libera combinazione di questi fra loro e con le diverse teorie e i diversi metodi utilizzati per studiarli [ISKO Italia 2007].

Connettere istituzioni

Limitazioni analoghe a quelle delle suddivisioni disciplinari, e di impedimento ancora maggiore sul piano pratico, si incontrano anche al livello delle istituzioni preposte a raccogliere i documenti, indicizzarli e permetterne l’accesso. Negli ultimi anni le comunità dei bibliotecari, degli archivisti e dei museologi hanno ri-

conosciuto che i loro mestieri (un termine che per inciso trovo ancora più efficace di “professioni” per non trascurarne il sapere apportato dall’esperienza) hanno molti punti in comune nel dover trattare risorse conoscitive, al di là dei diversi aspetti fisici con cui queste si presentano e delle ragioni istituzionali per le quali sono state raccolte: è ciò che ha portato ad accordi di collaborazione fra l’IFLA e le omologhe federazioni internazionali di archivisti e museologi, e in Italia alla creazione di MAB.

In una definizione più inclusiva della nozione di *documento* [Ridi 2010], pensato come qualsiasi supporto messo a disposizione per veicolare conoscenze, le “istituzioni della memoria” possono comprendere anche le raccolte di dipinti, in modo da ottenere l’accattivante sigla GLAM (*galleries, libraries, archives and museums*); i monumenti, non collocati insieme ai documenti prototipici essenzialmente per ragioni di spazio; le mostre, che differiscono dai musei solo nella durata (la scelta e disposizione dei padiglioni di Expo 2015 è una questione di organizzazione della conoscenza!); gli orti botanici e i giardini zoologici, pensati come musei di esemplari viventi, ognuno scelto per documentare l’intera specie a cui appartiene, con l’ausilio di ambientazione, segnaletica e pannelli illustrativi ecc.

L’esperienza di ricerca di informazioni mostra infatti che, analogamente all’interdisciplinarietà, sarebbe spesso necessaria anche una forte intermedialità. Chi cerchi di ricostruire la passata presenza di strumenti popolari ad ancia in Italia settentrionale trova di fatto informazioni rilevanti molto meno nei libri di etnomusicologia che negli archivi di polizia locale, nei musei di manufatti o negli affreschi di chiese sparse sul territorio [Gnoli 2010]. Ulteriori fonti di informazione possono apparire talmente inusuali da giungere del tutto inaspettate, prima che fosse venuto alla mente di cercare in quel tipo di supporti: così un personaggio conservato al Museo della marionetta di Torino, recando con sé una cornamusa i cui bordoni sono innestati in punti diversi dell’otre, può documentare la presenza di strumenti di origine locale e non, come si ritiene solitamente, provenienti da ambulanti abruzzesi; e il sipario ottocentesco del teatro comunale di Broni, nella scena di festa rurale che vi è raffigurata a tempera da Achille Savoia, può fornire indizi sulla presenza di strumenti a fiato nelle orchestre attive all’epoca nella zona.

Purtroppo un “catalogo unico” che consenta di cercare lo stesso soggetto in tutte queste fonti non è ancora disponibile, e c’è da temere che non lo sarà a lungo. Anche

in questo caso, comunque, un'organizzazione della conoscenza basata sui fenomeni prima che sulle istituzioni o sui supporti può cominciare a far avanzare in questa direzione. Un passo intermedio, già avviato in siti nazionali canadesi o scozzesi, consiste nell'indicizzazione del soggetto complessivo di intere collezioni di musei, archivi e biblioteche disponibili in uno stesso territorio [Dunsire 2005].

È chiaro che per rappresentare adeguatamente i contenuti di documenti di tal genere non tutti i KOS siano ugualmente adeguati: la Classificazione Dewey ad esempio ci obbligherebbe a stabilire che la marionetta del pastore natalizio appartenga alla disciplina della religione, oppure a quella del teatro o a quella dell'etnografia, e comunque non certo alla musicologia che invece interessa l'utente del nostro esempio. In questo senso i tesauri, come il Nuovo Soggettario BNCf o l'Art and Architecture Thesaurus, sono spesso meno vincolati ai contesti disciplinari rispetto alle classificazioni.

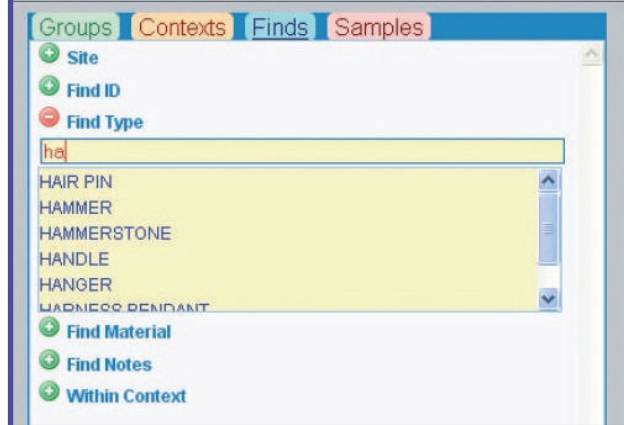
Appare comunque evidente l'opportunità di estendere la nozione di KOS anche a strumenti di origine diversa, come la sofisticata ontologia CIDOC-CRM sviluppata per l'indicizzazione di reperti museali. Le applicazioni ibride si osservano ormai con frequenza, ad esempio presso la Hypermedia research unit della University of South Wales che, in collaborazione con istituzioni regionali, applica un insieme di tesauri e classificazioni all'indicizzazione automatica di informazioni sui reperti archeologici presenti sul territorio inglese e gallese [Vlachidis et al. 2013].

Connettere KOS

Preso atto dell'esistenza di KOS diversi per contesto culturale, disciplinare, istituzionale, e per così dire nell'attesa del KOS generale e onnicomprensivo – che per qualcuno è concettualmente impossibile e per altri è almeno un'ideale direzione di sviluppo – ci troviamo per il momento a dover gestire un mondo di KOS plurali. A questo punto risulta evidente l'opportunità che anche i KOS stessi, e non solo i documenti e i dati con essi indicizzati, entrino a far parte del Web dei dati. Ciò richiede la rappresentazione dei KOS in formati standard e la loro messa a disposizione sulla Ragnatela globale.

I primi passi per la rappresentazione codificata di soggetti si possono far risalire al tempo di UNIMARC, che prevede campi specifici, in particolare nel blocco 6, per esprimere non solo la stringa di soggetto e il numero di classe attribuiti a un determinato documento, ma an-

FIGURE 1. Controlled term suggestion in the STAR Demonstrator.



Ricerca di termini usati nell'indicizzazione di documentazione archeologica dalla University of South Wales

che metadati specifici dei KOS quali il numero dell'edizione adottata o i componenti utilizzati per la costruzione di numeri di classe composti. Raramente tuttavia questi dettagli sono stati effettivamente impiegati e scambiati nei nostri OPAC.

Un primo sforzo in questa direzione all'interno dei nuovi linguaggi di rappresentazione per il web si è avuto con i metadati Dublin Core, che prevedono un elemento subject utilizzabile sia in forme libere che adottando un KOS esplicitamente indicato, quale uno schema di classificazione o un soggettario in una lingua specificata. Nel frattempo è nato un gruppo informale internazionale di esperti denominato NKOS, ossia appunto Networked KOS, che organizza annualmente un workshop in Europa e uno nel continente americano.

Il primo strumento espressamente concepito per rappresentare in RDF interi KOS bibliografici è stato però SKOS (Simple knowledge organization system): un nome alla lettera poco significativo, che induce qualcuno a credere che KOS si chiamino soltanto gli schemi codificati con esso. Al contrario, esiste un numero fin troppo grande di KOS ancora disponibili soltanto in forma di pesanti tomi a stampa, che nonostante la loro grande modernità concettuale (come nel caso della Bliss Classification 2) rischiano oggi di scomparire, vittime dell'imperativo "publish as RDF or perish".

D'altronde SKOS, nato sostanzialmente per rappresentare le relazioni standard dei tesauri, si rivela anche troppo *simple* non appena si cerchi di utilizzarlo per strutture comprendenti relazioni diverse, come le faccette, ritenute dalla teoria un elemento fondamen-

tale di qualsiasi KOS moderno. La registrazione di tali strutture di connessione concettuale è finora rimasta allo status di proposte di future estensioni di SKOS, per ora *postponed*, forse per effetto dell'egemonia americana che prende in considerazione soltanto KOS ipertradizionali non veramente faccettati, quali LCSH, LCC e CDD: in Europa avremmo di meglio, ma non possiamo ancora SKOSificarlo adeguatamente.

Una via alternativa suggerita da chi ha notato queste limitazioni [Zeng et al. 2010] è quella di utilizzare invece OWL, altro linguaggio di rappresentazione della conoscenza nato nel contesto delle ontologie (dove già ne esistevano antecedenti come DAML+OIL), che proprio per questo consente di definire e utilizzare qualsiasi tipo di relazione. Si tratta però di un percorso ancora da sviluppare, definendo e condividendo gli equivalenti RDF di relazioni concettuali da tempo identificate dalla ricerca in organizzazione della conoscenza, quali gli ausiliari comuni, le faccette, le relazioni di fase o la dipendenza esistenziale, ma che non sono ancora state formalizzate per essere adeguatamente sfruttate nel nuovo contesto dei linked data. La ragione, ci sembra, come al solito non sta in una loro eccessiva difficoltà o poca utilità, ma nella scarsità di investimenti negli aspetti concettuali rispetto a quelli tecnologici, più attraenti e alla moda con le loro prestazioni eclatanti ma relativamente superficiali dal punto di vista dell'organizzazione della conoscenza.

Una volta debitamente rappresentati come linked data, i KOS possono essere messi in relazione fra loro, oltre che con i dati e i documenti con essi indicizzati. Questa operazione di *mapping* deve evidentemente tener conto dei limiti concettuali di cui abbiamo parlato sopra, accontentandosi talvolta di *broad match* anziché *exact match*. Ma proprio la consapevolezza di tale approssimazione e la sua dichiarazione formale possono essere una soluzione. Promettente in questo senso è la recente pubblicazione del nuovo standard ISO per l'interoperabilità fra tesauri e altri vocabolari controllati, che prevede anche l'equivalenza approssimativa fra termini e l'equivalenza fra un termine di un KOS e l'unione di altri due di un altro KOS (nell'esempio pubblicato, inland waterways EQ rivers | canals) [Dextre Clarke 2012].

Conclusioni

La nostra rapida rassegna ha mostrato come la connessione dei contenuti, lungi dall'essere risolta dalla connessione tecnica dei dati, è possibile soltanto costruendo

sulle spalle di quest'ultima una effettiva interoperabilità dei soggetti. Gli strumenti che la renderebbero possibile, al di là della superficiale varietà di gerghi provenienti da comunità diverse che il digitale porta a convergere, non possono che essere dei *knowledge organization system*. Questi spesso soffrono di limitazioni che li rendono in realtà meno interculturali, interdisciplinari, intermediali e tecnicamente interoperabili di quanto sarebbe richiesto da una completa connettibilità concettuale. Quest'ultima è forse un'utopia, ma ogni passo che si compie in questa direzione promette ricadute potenti: la strada in questa direzione ci appare oggi più chiara che in passato.

Ringraziamenti

Vincenzo Maltese (Università di Trento) e Riccardo Ridi (Università Ca' Foscari di Venezia) mi hanno fornito utili commenti a una versione precedente del testo.

BIBLIOGRAFIA

- GIULIO BARSANTI, *La scala, la mappa, l'albero: immagini e classificazioni della natura fra Sei- e Ottocento*, Firenze, Sansoni, 1992.
- CLARE BEGHTOL, *Semantic validity: concepts of warrant in bibliographic classification systems*, "Library resources and technical services", 30 (1986), n. 2, p. 109-125.
- C. BEGHTOL, "Facets" as interdisciplinary undiscovered public knowledge: S.R. Ranganathan in India and L. Guttman in Israel, "Journal of documentation", 51 (1995), n. 3, p. 194-224.
- MARIA TERESA BIAGETTI, *Le ontologie*, "AIDA informazioni", 28 (2010), n. 1-2.
- ROY DAVIES, *Generating new knowledge by retrieving information*, "Journal of documentation", 46 (1990), n. 4, p. 368-372.
- STELLA DEXTRE CLARKE, *ISO 25964: a standard in support of KOS interoperability*, in *Facets of knowledge organization: proceedings of the ISKO UK Second Biennial Conference, 4th- 5th July 2011, London*, a cura di A. Gilchrist - J. Vernau, Bingley, Emerald, 2012, p. 129-134.
- GORDON DUNSIRE, *Conspectus and the Scottish Collections Network: landscaping the Scottish common information environment*, "Signum", 306, <http://www.varastokirjasto.fi/kokoelmakartta/julkaisut/artikkelit/Signum306_Dunsire.pdf>.
- ANTONY CHARLES FOSKETT, *Il soggetto*, traduzione di L. Bultrini, Milano, Editrice Bibliografica, 2001.
- IVANO FOSSATI, *Labile*, in *Macramé*, CD audio, Columbia, 1996.
- FAUSTO GIUNCHIGLIA - BISWANATH DUTTA - VINCENZO MALTESE, *From knowledge organization to knowledge representation*, "Knowledge organization", 41 (2014), n. 1, p. 44-56.

- CLAUDIO GNOLI, *Classification transcends library business*, "Knowledge organization", 37 (2010), n. 3, p. 223-229.
- C. GNOLI - RICCARDO RIDI - GIULIA VISINTIN, *Di che parla questo catalogo?*, "Biblioteche oggi", 22 (2004), n. 8, p. 23-29.
- HEATHER HEDDEN, *Taxonomies vs. thesauri*, "The accidental taxonomist" [blog], January 28, 2014, <<http://accidental-taxonomist.blogspot.it>>.
- BIRGER HJØRLAND, *Information organization=knowledge organization*, in *Categories, contexts and relations in knowledge organization: proceedings of the Twelfth International ISKO Conference, Mysore, India, August 6-9, 2012*, a cura di A. Neelameghan - K.S. Raghavan, Würzburg, Ergon, 2012.
- ISKO ITALIA, *The León Manifesto*, <<http://www.iskoi.org/ilc/leon.php>>, anche in "Knowledge organization", 34 (2007), n. 1, p. 6-8.
- MANUEL LIMA, *The book of trees: visualizing branches of knowledge*, New York, Princeton Architectural Press, 2014.
- CARLO REVELLI, *Soggettazione a faccette*, in *Le dimensioni dell'informazione*, Documenti ISKO Italia, 2004, <<http://www.iskoi.org/doc/dimensioni1.htm>>.
- RICCARDO RIDI, *Il mondo dei documenti: cosa sono, come valutarli e organizzarli*, Roma, Laterza, 2010.
- MICHELE SANTORO, *La disarmonia prestabilita: per un approccio ibrido alla conoscenza e ai suoi supporti*, convegno "La biblioteca ibrida", in "Biblioteche oggi", 20 (2002), n. 5, p. 46-57; commenti di C. Gnoli e di G. Visintin in "Biblioteche oggi", 21 (2003), n. 1.
- FRANCESCA SEVERINO, *What thesaurus to define EU/ACP relations?: analysis of the term development in the thesauri of the EU and other international organizations*, Documenti ISKO Italia, 2006, <<http://www.iskoi.org/doc/development.htm>>, versione rivista in "European journal of developmental research", 19 (2007), n. 2, p. 251-327.
- RICHARD P. SMIRAGLIA, *Be careful what you wish for: FRBR, some lacunae, a review*, "Cataloging & classification quarterly", 50 (2012), p. 360-368.
- RICK SZOSTAK, *Classification, interdisciplinarity, and the study of science*, "Journal of documentation", 64 (2008), n. 3, p. 319-332.
- ANDREAS VLACHIDIS - CERI BINDING - KEITH MAY - DOUGLAS TUDHOPE, *Automatic metadata generation in an archaeological digital library: semantic annotation of grey literature*, in *Computational linguistics, a cura di A. Przepikowski - M. Piasecki - K. Jassem - P. Fuglewicz*, Berlin-Heidelberg, Springer, 2013, p. 187-202.
- DAVID WEINBERGER, *Elogio del disordine: le nuove regole del mondo digitale*, Milano, BUR Rizzoli, 2010.
- SCOTT B. WEINGART, *From trees to webs: uprooting knowledge through visualization*, in A. Slavic - A. Akdag Salah - S. Davies (a cura di), *Classification and visualization: interfaces to knowledge*, Würzburg, Ergon, 2013, p. 53-67.
- ROBIN JAMES WILSON, *Introduzione alla teoria dei grafi*, Roma, Cremonese, 1978.
- MARCIA LEI ZENG - MICHAEL PANZER - ATHENA SALABA, *Expressing classification schemes with OWL 2: exploring issues and opportunities based on experiments using OWL 2 for three classification schemes*, in *Paradigms and conceptual systems in knowledge organization: proceedings of the Eleventh International ISKO Conference, 23-26 February 2010, Rome, Italy*, a cura di C. Gnoli - F. Maz-zocchi, Würzburg, Ergon, 2010, p. 356-362.

DOI: 10.3302/0392-8586-201405-009-1

ABSTRACT

The technological connection of data may be useless, if we do not make their meanings connected as well in some way. To this purpose, what in traditional libraries is called subject indexing, and in the current broadening of contexts, more in general, knowledge organization, is coming back to our attention. Indeed, several kinds of knowledge organization systems (KOS) enable us to connect information according to its contents, irrespective of the variety of formats, carriers, location. However, this is only possible if KOS are interoperable, capable of overcoming information boundaries at four layers at least: a) conceptual, that is depending on different theoretical and cultural approaches; b) disciplinary, depending on the traditions and languages of different social and academic communities; c) institutional, depending on the different policies of libraries, archives, and museums holding knowledge sources, including those hidden in such unusual documents as a theatre curtain or a puppet; d) formal, depending on representation languages (MARC, SKOS, OWL etc.) enabling to leverage KOS in the global network of linked data.