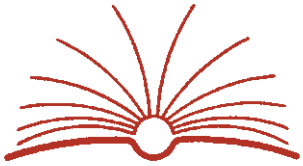

Knowledge organization and classification in international information retrieval

Nancy J. Williamson, Clare Beghtol (eds.), Binghamton, NY, The Haworth information Press, 2003, ISBN 0-7890-2354-7
Pubblicato anche in "Cataloging and classification quarterly", 37 (2003), 1/2

Il volume, pubblicato dalla Haworth Press come consuetudine anche in forma di fascicolo di rivista, tratta il tema del recupero dell'informazione in ambito internazionale, attraverso sistemi di organizzazione della conoscenza e schemi di classificazione.

La trattazione è affidata a quattordici articoli di diversa ampiezza, suddivisi in quattro sezioni che affrontano il tema nei suoi aspetti più generali e poi via via più specifici, fino ad arrivare alla descrizione di alcune esperienze modello. L'opera mette a confronto autori di diverse nazionalità ed esperienza professionale, proprio per allargare il numero di linguaggi, domini della conoscenza e dell'informazione, risorse informative e differenti approcci messi in atto per risolvere il problema di quella che i curatori chiamano "globalization of information and knowledge". Il tema centrale dell'opera è, infatti, quello dell'organizzazione di sistemi di information retrieval e di proporre soluzioni applicabili ad ambienti multilingue, con schemi di classificazione che permettano il passaggio da una lingua all'altra, con la relativa difficoltà di mappare i diversi linguaggi informativi.

La prima parte, "General bibliographic systems", riguarda il futuro dei sistemi generali di classificazione, e mostra come essi possano essere adattati alle necessità di specifici campi disciplinari, come la loro struttura concettuale possa essere comparata ed eventualmente modificata e quali, infine, possano essere le difficoltà di traduzione dall'originario contesto linguistico-culturale a una diversa lingua e a un'altra cultura. Il primo intervento (*The future of general classification*) offre una panoramica sulla diffusione dell'informazione, in particolare nel World Wide Web, soffermandosi sul tema chiave dell'interoperabilità, che viene definito su tre livelli: tecnico, contenutistico e organizzativo. Viene analizzato in particolare il secondo aspetto e le due possibili soluzioni al problema dell'interoperabilità semantica nell'accesso per soggetto: l'uso dello stesso (o quasi) linguaggio di indicizzazione o la costruzione di un linguaggio intermedio, di scambio (*switching* o *cross-language*). Il problema, sottolineato, è quello che nel passaggio tra due lingue si possano allargare o restringere i campi semantici di alcuni termini, che quindi assumono significati più o meno diversi, più o meno ampi, in lingue differenti, L'esempio portato (p. 8), quello dei temini inglesi *tree*, *wood*, *woods*, *forest*, è assai significativo per mostrare come queste parole non abbiano un corrispettivo "uno a uno" nelle altre cinque lingue prese in esame, dato che si va da due termini del danese a cinque dello spagnolo. Nella tabella che mette a confronto i termini, un refuso trasforma l'atteso "bosco" in un inaspettato "basco", cosa che capita anche in spagnolo: *basque* per *bosque*. Questioni legate alla traduzione di classificazioni sono presentate anche in un successivo contributo, che mostra la capacità di adattamento di un sistema di classificazione generale (la CDD) a uno specifico campo di studi (quelli sul femminismo, assai interessante per l'ottica interdisciplinare). Il progetto descritto ha ottenuto due risultati: innanzitutto un'espansione della CDD per l'organizzazione di collezioni con una certa enfasi sugli studi femministi e in aree vicine, e poi la sperimentazione di un metodo (consistente nell'identificazione di carenze, sviluppo e implementazione di un piano di soluzione che preveda la

compatibilità con il resto dello schema, la credibilità dei nuovi sviluppi e la facilità della loro applicazione, e infine di una fase di test dell'efficacia delle scelte proposte) che potrebbe servire da modello per ulteriori analoghe espansioni.

Tutta la seconda sezione, "Information organization in knowledge resources", dovrebbe essere utilizzata come dispensa nei corsi per bibliotecari che soffrono della Sindrome della foca monaca. Le biblioteche continueranno a esistere e i bibliotecari non si estingueranno, almeno non nel prossimo futuro… Tutti gli autori sono infatti concordi nel riconoscere grande efficacia sia a sistemi pre- che post-coordinati, ma anche sulla necessità di rendere ancora più solide le basi concettuali nei servizi di accesso semantico all'informazione, che si potranno poi applicare anche in contesti allargati rispetto agli attuali cataloghi di biblioteca.

La terza parte, "Linguistics, terminology and natural language processing", è una rassegna di teorie e tecniche che mettono in rilievo il ruolo essenziale della linguistica nell'accesso, nell'estrazione e nella disseminazione dell'informazione. Il problema fondamentale emerge già nel primo contributo della sezione: "Most user can read documents written in a foreign language but have difficulty finding the appropriate words to express their query in an efficient way" (p. 140). In questo intervento, infatti, vengono passate in rassegna, poco più che accennate a dire il vero, differenti tecniche utilizzate nel CLIR, Cross-Language Information Retrieval: traduzione automatica, indicizzazione automatica latente, costruzione automatica di thesauri *cross-language*, riformulazione delle query impostate dall'utente con sistemi automatici secondo criteri morfo-sintattici ecc. Tra gli altri articoli mi ha colpito, e lasciato a bocca aperta, quello che presenta i risultati di un'indagine sul corretto riconoscimento dei nomi, le etichette, assegnati ai vari campi di un record bibliografico. Prima parte dell'esperienza è stata l'identificazione di diciassette campi e delle relative etichette, ottenuta attraverso la mappatura di nove diversi schemi di metadati tra i più conosciuti (EAD, Dublin Core, GILS, TEI, VRA, CIMI, CSDGM, ONIX, MARC). Nell'esperimento presentato (p. 174-185), diciannove "cavie" (dodici con esperienza formalizzata in tecniche di ricerca informativa, sette senza) dovevano associare diciassette etichette di elementi alle rispettive definizioni (ad esempio: *edition = information on a work's version*). Il risultato, estremamente sconfortante, è stato che soltanto quattro persone hanno completato correttamente il test, e solo tre elementi sono stati associati correttamente da tutti i partecipanti alle rispettive etichette (*physical format, language, date & time period*). La cosa inquietante è che tra quelli rimasti fuori ci fosse addirittura… *title*!

Nella quarta e ultima parte, "Knowledge in the world and the world of knowledge", vengono infine messi a confronto modelli di KM (Knowledge Management) di differenti matrici culturali e l'effetto, in un medesimo dominio della conoscenza, dell'uso di diversi sistemi di classificazione e di organizzazione. Dopo una brevissima analisi bibliometrica (tesa a verificare quali siano gli studiosi maggiormente citati in ambito KM, identificati da Ikujiro Nonaka e Thomas H. Davenport), l'autore di *Perspectives on managing knowledge in organizations* passa alla verifica della sua ipotesi, se cioè le scuole di pensiero che questi due autori rappresentano possano in qualche modo riflettere "usi e costumi delle loro culture nazionali" (Giappone e USA). L'interessante esame, dall'esito almeno in parte prevedibile, dimostra che l'assunto iniziale è vero. L'approccio di Nonaka, attraverso la sua definizione della conoscenza tacita e della sua continua trasformazione in conoscenza esplicita e viceversa ("embracing opposites", p. 218), riconosce l'unicità di corpo e mente, mantenendosi pienamente nell'alveo del pensiero tradizionale giapponese; viceversa, le idee di cui si fa portatore Davenport sono nate da un confronto serrato con i dirigenti di alcune tra le maggiori aziende statunitensi sul loro approccio all'informazione. D'altra parte, certi concetti proposti parrebbero in opposizione con i modelli culturali nazionali ma questo, secondo l'autore, sarebbe un modo per bilanciare o compensare quelle tendenze culturali che potrebbero impedire la creazione e il trasferimento della conoscenza (ad esempio, l'avversione della cultura giapponese al caos e alla crisi).

In conclusione, mi pare che questo volume soffra degli stessi difetti di altri del medesimo editore: il fatto che nasca come fascicolo di una rivista, infatti, mette assieme contributi assai dissimili tra loro per valore e tematiche trattate (alcuni di essi solo genericamente riconducibili al titolo e al filo conduttore dell'opera) e per grado di specificità e di approfondimento. A ciò, che potrebbe non essere necessariamente un grosso difetto, si aggiunga che manca (all'infuori di una presentazione di una pagina e mezzo, su un totale di 244) uno strumento di coesione dell'opera che possa guidare il lettore all'interno di un percorso organico di comprensione. In alcuni contributi (p. 13, 65, 81, 199, 221…), poi, l'apparato bibliografico mi pare un po' datato, non trovando citazioni successive al 2001 (il volume è stato pubblicato nel 2003!).

Ben costruito, invece, l'indice analitico, per nomi e parole citati, che comprende anche rinvii decisamente insoliti ma utili, del tipo *see also*. Curiosamente, e non sono riuscito a spiegarmene la ragione, gli autori dei contributi proposti, e non solo gli autori da essi citati, compaiono in tale indice, nonostante esista anche un sommario completo dell'opera che comprende tutti gli articoli e i relativi autori.

Andrea Marchitelli

Roma
andreamarchitelli@hotmail.com