

L'indicizzazione semantica al tempo del web

Laura Ballestra

A colloquio con Vanda Broughton, una delle maggiori esperte di classificazioni a faccette

Biblioteca "Mario Rostoni"
Università Carlo Cattaneo LIUC
lballestra@liuc.it

La Classificazione Bliss ha una storia molto lunga, che comincia nel 1940 e giunge fino alla revisione iniziata nel 1967. Dal 1977 lei ne è la curatrice, insieme a Jack Mills.

Qual è stata l'evoluzione di questa classificazione negli ultimi trenta anni e quali sono le attuali prospettive?

La Bliss Classification (BC) è cambiata molto in questo periodo e a partire dalla pubblicazione dell'*Introduction and auxiliary schedules*¹ del 1977. All'inizio c'era della tensione tra il desiderio di implementare la teoria che era stata sviluppata durante il ventennio precedente e il bisogno di soddisfare le attese degli utenti che desideravano che lo schema non cambiasse troppo.

I piani originari di revisione prevedevano una classificazione piuttosto ampia, comparabile come dimensioni alla Dewey, in cui i cambiamenti più rilevanti avrebbero dovuto riguardare la modernizzazione del vocabolario e la razionalizzazione dell'ordine di citazione. Al contrario, quando ci si rese conto che poche biblioteche avrebbero operato la conversione dallo schema BC1 alla BC2 (le biblioteche che usano la BC2 per lo più sono nuovi utenti della classificazione) la revisione diventò più radicale, con l'aggiunta per intero di vari vocabolari specialistici e una maggiore attenzione alle variazioni nell'ordine di citazione. La nuova

terminologia soddisfa meglio le esigenze di collezioni speciali ed è più adatta alla natura complessa dell'informazione digitale.

Ci può dire qualcosa del team di lavoro della BC2? Come lavorate?

La metodologia di revisione è cambiata alquanto negli anni. Fino agli anni Novanta i fondi assegnati originariamente per la revisione e i fondi di ricerca che sono arrivati in seguito, insieme alle sponsorizzazioni, sono stati impiegati per pagare il lavoro di un ricercatore a tempo pieno (la sottoscritta) e Jack Mills. Ciò non significa che tutto il lavoro sia stato fatto solo da noi due. Anche i membri della Bliss Classification Association (BCA), e in particolar modo il Classification Resource Group (CRG), hanno contribuito alla redazione delle tavole. Questo ha fatto sì che molte classi fossero sviluppate da altre persone piuttosto che dai curatori. Anche le tavole realizzate in prima persona da Mills e da me sono state commentate da un ampio gruppo di persone interessate: tutte sono state discusse dal CRG, e le bozze sono state sottoposte ad un gruppo significativo di utenti della BCA e ad altre persone che hanno fornito commenti e feedback preziosi. Questo continua ad essere il modello di lavoro, anche adesso che dipendiamo da contributi volontari. La collaborazione di alcune biblioteche del Cambridge College (molte delle quali adottaro-



Vanda Broughton, MA DipLib, è *lecturer* presso la School of Library, Archive and Information Studies dell'University College London (v.broughton@ucl.ac.uk). Insegna nei corsi *Cataloguing and classification 1*, *Information sources*, *Cataloguing and classification 2*. Dal 1977 è curatrice della BC2. È anche curatrice associata dell'Universal Decimal Classification, coordinatrice della sezione inglese di ISKO, segretario onorario della Bliss Classification Association.

no la BC2 negli anni Novanta) è stata molto rilevante, tanto che un cospicuo numero di bozze di nuove tavole è stato sviluppato proprio all'interno di quell'università.

Nel prossimo futuro speriamo di poter pubblicare le bozze delle tavole sul sito della BCA in modo che una comunità più ampia possa partecipare alla discussione. Per agevolare il lavoro intellettuale sulla classificazione, la produ-

zione fisica delle tavole è gestita da un insieme di programmi che generano la versione finale e l'indice alfabetico.

Di recente abbiamo acquisito nuovi programmi che ci consentono di derivare anche un thesaurus, e speriamo che questo possa attrarre una maggiore varietà di utenti, specialmente quelli che mantengono collezioni di risorse digitali.

Pensa che la BC2 possa essere tradotta in altre lingue?

In teoria, una classificazione è il tipo più semplice di vocabolario controllato da tradurre in un'altra lingua naturale, dato che i concetti sono rappresentati attraverso delle notazioni.

Questo significa che le vedette delle classi non hanno bisogno di essere particolarmente brevi, perché non sono usate nell'indicizzazione, e qualunque parte di testo che definisce o spiega una classe può essere inserita nella descrizione della classe stessa.

In pratica la situazione non è proprio così semplice, perché può non esserci esatta corrispondenza tra strutture concettuali dei soggetti in lingue differenti; per esempio in inglese non c'è una parola che indichi "slugs" + "snails", mentre esiste in tedesco e in olandese; quindi la gerarchia per i molluschi è diversa in lingue diverse, anche quando queste hanno origini comuni.

Inoltre alcuni dei problemi che caratterizzano i thesauri multilingue caratterizzano anche le classificazioni, ma in generale la natura concettuale della classificazione evita gli altri problemi associati con i linguaggi di indicizzazione.

Quindi non pare esserci alcun motivo per cui la BC2 non potrebbe essere tradotta, così come è accaduto per la Dewey e la UDC e per molti altri differenti linguaggi.

Possiamo dire che la BC2 è uno schema di classificazione comple-

tamente a faccette. Quali problemi si incontrano nel definire la struttura quando si lavora in discipline così diverse?

La metodologia standard utilizzata nella BC2 è, forse in maniera sorprendente, trasferibile e abbastanza ben utilizzabile in un ampio spettro di discipline. Per uno schema di classificazione universale l'ordine di citazione e i principi strutturali generali devono essere applicati in maniera coerente, altrimenti possono crearsi troppe situazioni imprevedibili e potenzialmente conflittuali, così che in una certa misura le necessità specifiche poste da singole discipline hanno dovuto essere considerate secondarie rispetto alla struttura generale della classificazione. In pratica il numero delle alternative consentite della BC2 permette di soddisfare le richieste poste da differenti tipi di collezioni e di utenti. Noi partiamo sempre utilizzando il normale insieme delle categorie del CRG (cose, tipi di cose, parti di cose, materia/materiali, proprietà, processi, attività, prodotti, prodotti intermedi, agenti/strumenti, spazio, tempo) e applicando l'ordine di citazione standard, ma spesso questo modo di procedere richiede alcune modifiche per specifici aspetti disciplinari.

Quasi sempre questa necessità di modifiche emerge abbastanza chiaramente dall'analisi della letteratura disciplinare e relazionandosi con gli specialisti.

Quello che è inevitabile è la necessità di impadronirsi a fondo della terminologia di una disciplina. Questo è l'aspetto che richiede il maggior investimento di tempo durante il lavoro di revisione, specialmente per argomenti tecnici nei quali le relazioni tra i termini possono non essere così ben evidenti a prima vista.

Questo è stato certo il caso della chimica e della matematica, una disciplina particolarmente astratta e difficile.

Spesso è anche necessario l'uso di qualche "nuova" categoria, non inclusa nell'elenco standard. Le forme e i generi sono esempi ben conosciuti impiegati nel caso delle discipline artistiche.

Qualunque sia l'argomento, al centro dell'analisi per faccette c'è un esame estremamente dettagliato e preciso della terminologia, e una altrettanto chiara specificazione delle relazioni tra i termini.

Quante biblioteche stanno usando la BC2?

Attualmente ci sono circa cinquanta biblioteche che usano la Bliss Classification, e quasi tutte sono utilizzatrici della BC2. Nuove biblioteche continuano ad aggiungersi, ma l'uso potenzialmente maggiore che la BC2 potrebbe avere è come strumento di indicizzazione più generale (da intendersi come opposto a un sistema utilizzato per l'organizzazione fisica di risorse documentali). Questa convinzione è quella che ci ha spinto a cominciare l'esplorazione della possibilità di derivare thesauri direttamente dalla classificazione stessa.

Lei è un'esperta sia di thesauri che di schemi di classificazione. Qual è la relazione a suo parere tra questi due strumenti?

Una volta il thesaurus e lo schema di classificazione sarebbero stati visti come due strumenti di tipo abbastanza diverso, ma il lavoro di Jean Aitchinson con il *Thesaurusface²* e i successivi thesauri a faccette hanno mostrato come i principi dell'analisi a faccette potrebbero essere usati per costruire entrambi i tipi di vocabolari. Infatti l'utilizzo dell'analisi a faccette per creare una struttura sistematica è un metodo molto utile per identificare le relazioni tra i termini che possono essere necessari in un thesaurus. Dopo il *Thesaurusfacet* è diventato abbastanza comune pubblicare un thesaurus in due parti: una struttura sistematica

ca o classificata e una lista alfabetica dei termini con le relazioni thesaurali. La metodologia di Jean è stata seguita nel lavoro attuale per il thesaurus della BC2, nel quale noi abbiamo potuto utilizzare un vocabolario a faccette come base sia per la classificazione definitiva, sia per l'indice alfabetico della classificazione che per un vero e proprio thesaurus. Questo è stato ottenuto attraverso la codifica del vocabolario sorgente, in modo tale che i tre tipi di presentazione possono essere generati da programmi in modo quasi totalmente automatico.

Ciò dimostra, a mio parere, che le differenti presentazioni non sono altro che un diverso modo di esporre la stessa terminologia. Quello che uno scopre è che etichette come thesaurus, tassonomia, ontologia... sono attualmente usate in modo intercambiabile e che in particolare i non specialisti in biblioteconomia e scienze dell'informazione non vedono la necessità di fare queste distinzioni tra strumenti di indicizzazione.

Nel Regno Unito molti siti web stanno usando thesauri come vocabolari controllati per l'indicizzazione di pagine web. Pensa che il thesaurus sia una buona applicazione per i metadati semantici dei siti web?

Penso che il thesaurus sia davvero eccellente come strumento per assegnare metadati semantici, dal momento che la struttura del thesaurus consente una gestione dei metadati più sofisticata.

Per esempio, le relazioni dei thesauri forniscono un modo facile per navigare nelle collezioni di risorse documentali e consentono di esplorare i materiali collegati allargando o restringendo la ricerca, attraverso BT e NT, oppure cercando anche con i termini collegati. Esiste anche un uso potenziale della presentazione sistematica del thesaurus come una struttura di navigazione, particolarmente per le ri-

sorse digitali, contesto nel quale il sistema dei collegamenti ipertestuali (link) sembra idealmente progettato proprio per questo.

La natura di linguaggio controllato del vocabolario di un thesaurus dovrebbe inoltre migliorare il recupero in generale, e un linguaggio controllato può essere usato dietro un'interfaccia per formulare o modificare l'interrogazione dell'utente senza che questi debba accorgersene. È interessante notare che molti studi recenti sulla generazione automatica di metadati hanno mostrato come l'estrazione automatica dei termini dai testi e il successivo riscontro e collegamento (mappatura) di questi con un thesaurus fornisce un modo poco costoso di assegnare metadati di qualità senza l'intervento umano.

La Biblioteca nazionale centrale di Firenze sta adottando un thesaurus a faccette per lo sviluppo del Nuovo soggettario. Cosa pensa della possibilità di usare un thesaurus generale come fonte di vocabolario nella costruzione di stringhe di soggetto?

Quando insegno agli studenti come costruire un thesaurus, l'esercizio termina sempre impiegando il thesaurus prodotto per indicizzare alcuni dei documenti da cui all'inizio sono stati ricavati i vocaboli che hanno costituito la base terminologica per lo sviluppo del thesaurus. L'operazione viene completata mostrando come questi termini di indicizzazione possano essere usati per costruire stringhe di soggetto, applicando al loro insieme l'ordine di citazione standard.

Il procedimento funziona molto bene, e sebbene le stringhe risultanti non siano esattamente uguali a quelle di un linguaggio preordinato, come quelle del soggettario della Library of Congress, sono comunque molto regolari nella struttura e io credo più facili da capire per gli utenti.

Il progetto FAST della Library of

La Classificazione Bliss e BC2

Il Bibliographic Classification System, meglio noto come Bliss Classification, nasce dal lavoro di Henry Bliss (1870-1955), che elaborò il suo schema di classificazione a partire dai primi del Novecento. Bibliotecario presso il City College of New York, Bliss era critico verso la Classificazione decimale Dewey, il sistema più impiegato all'epoca, e nel 1908 impiegò la nuova classificazione da lui ideata per riclassificare l'intera collezione della Biblioteca del City College.

Tra le sue pubblicazioni più note figura *A sistem of bibliographic classification* del 1935.

L'intero schema della classificazione da lui realizzato venne pubblicato in quattro volumi dal 1940 al 1953 e impiegato da oltre ottanta biblioteche, soprattutto in Gran Bretagna e nei paesi del Commonwealth.

Agli inizi degli anni Settanta lo schema subì una revisione radicale che portò alla creazione di un nuovo sistema, la BC2, che dell'originario impiega solo la struttura principale. I più importanti artefici della BC2 sono stati Jack Mill e Vanda Broughton, ma fondamentale è risultata la riflessione maturata in seno al Classification Resource Group (CRG).

La BC2 prevede che tutte le classi, ed eventualmente le sottoclassi, siano strutturate a faccette; un ordine di citazione unitario e coerente è presente in tutte le classi; l'ordine di elencazione mantiene sempre il generale prima del particolare; la notazione è sintetica e completamente strutturata per faccette; sono forniti indici alfabetici dettagliati per tutte le classi.

La classificazione venne pubblicata a partire dal 1977 da Butterworth, ora è edita da KG Saur.

Dei ventitré volumi previsti ne sono stati editati ad oggi quattordici.

La Bliss Classification Association (<http://www.blissclassification.org.uk/>), che è detentrica del copyright, cura lo sviluppo e gli aggiornamenti.

Il "Bliss Classification Bulletin" è lo strumento attraverso cui i membri dell'associazione e gli utilizzatori della BC2 vengono a conoscenza dell'evoluzione e dell'aggiornamento delle tavole.

Congress, che ha come obiettivo quello di razionalizzare e semplificare le liste di soggetto della Library of Congress per un uso in ambiente digitale, impiega stringhe di soggetto con una struttura simile, il che ci suggerisce che anche loro la pensano in questo modo.

Mantenere un thesaurus o adottare uno schema di classificazione ha un costo. Pensa che l'attuale evoluzione di Internet, tra Web 2.0, ontologie e tassonomie, abbia determinato una maggiore attenzione al problema dell'indicizzazione dell'informazione?

Sì, penso che oggi ci sia molta più attenzione ad un approccio ben strutturato per il recupero di informazione sul web, e che molti informatici sono oggi davvero più

consapevoli degli aspetti teoretici connessi ad una buona architettura dell'informazione, a scapito della prospettiva di ritenere che un processore molto veloce possa da sé risolvere tutti i problemi.

Dieci anni fa avrei potuto semplicemente dire che gli informatici avevano molto da imparare dagli specialisti di biblioteconomia e scienze dell'informazione, ma oggi penso che loro abbiano davvero intrapreso questo processo, e che ci sia più consapevolezza del rilievo di tecniche come l'analisi per faccette e dei diversi modi per strutturare i thesauri al fine di migliorare l'organizzazione e il recupero.

C'è ancora, certamente, un grande investimento della ricerca volto a trovare modi di generare automaticamente thesauri e ontologie, ma

i risultati sono quasi tutti molto meno sofisticati e non così efficaci per indicizzare e recuperare documenti rispetto a quelli offerti dagli strumenti costruiti dall'uomo.

Come pensa che i bibliotecari possano presentare agli informatici le loro conoscenze sull'indicizzazione per favorire l'uso di classificazioni e thesauri nel web?

Come ho già detto penso che oggi gli informatici siano più consapevoli delle teorie LIS di quanto non fossero in passato, ma noi dobbiamo continuare a ricercare, a comunicare e a scrivere su quello che conosciamo, così che un corpus di conoscenze sia disponibile per tutti.

Credo che una delle ragioni per cui la teoria sulle classificazioni ha avuto una presa così lenta presso gli specialisti IT sia da ricercarsi nella disseminazione molto limitata dei risultati della ricerca, e che tutto si sia realizzato all'interno di un gruppo molto ristretto.

Questo vale sicuramente per il lavoro del CRG, che è raccolto in un numero molto limitato di pubblicazioni ed ha richiesto molto tempo per diventare oggetto di attenzione da parte di una comunità più ampia.

Mentre i ricercatori oggi possono avere maggior familiarità con le teorie LIS, c'è ancora bisogno che i bibliotecari parlino direttamente per sé, che parlino delle proprie capacità nei contesti lavorativi, specie là dove i tecnici informatici tendono a considerarli come non competenti in materia di recupero dell'informazione. È necessario mantenere un ponte tra aspetti teorici, pratici e tecnici nella catalogazione e dell'indicizzazione, e assicurare che tutti restino in equilibrio.

È molto più probabile rispetto al passato che i nuovi professionisti LIS acquisiscano competenze tecniche come parte della loro formazione professionale, ma dobbiamo continuare ad insistere perché non

si cerchi di evitare questi aspetti più "difficili" del lavoro, perché questo significherebbe essere messi da parte. Solo se capiremo meglio la tecnologia informatica sapremo davvero andare d'accordo con gli informatici e parlare loro con la stessa lingua.

Ha visto molti cambiamenti tecnologici negli ultimi trent'anni. Qual è dal punto di vista dell'attività di indicizzazione semantica il cambiamento più significativo?

Credo che il cambiamento più significativo stia nel passaggio da

collezioni cartacee a digitali e ibride. Questo ha coinvolto tutte le aree della pratica professionale, e ovviamente anche l'indicizzazione. Quando, terminati gli studi, ho cominciato a lavorare come ricercatrice negli anni Settanta, Internet, materiali digitali e cataloghi elettronici non esistevano, e c'erano solo un numero limitato di servizi online come Medline, l'indice alla letteratura medica.

I primi computer cominciavano a essere introdotti nelle biblioteche, e come studenti di biblioteconomia ci veniva insegnato cosa *avrebbero po-*

Pubblicazioni di Vanda Broughton

Bliss Bibliographic Classification 2nd edition Class W Fine arts, (with Jack Mills), Munich, Saur, 2006.

Building a faceted classification for the humanities: principles and procedures, (with Aida Slavic), "Journal of Documentation", 63 (2007), 5, p. 727-754; <http://dlist.sir.arizona.edu/1976/01/Broughton_Slavic_jdoc2007_preprint.pdf>.

Classification and subject organization and retrieval, in BOWMAN (ed.), *British librarianship and information work 2001-2005*, Aldershot, Ashgate, 2007, p. 467-488.

Classification and subject organization and retrieval, in BOWMAN (ed.), *British librarianship and information work 1991-2000*, Aldershot, Ashgate, 2006, p. 494-516.

Una classificazione per il XXI secolo. Principi e struttura della Classificazione bibliografica Bliss, "AIB Web contributi", 2004, <<http://www.aib.it/aib/contr/broughton1.htm>> (traduzione di Claudio Gnoli dall'originale inglese dell'articolo in croato *Klasifikacija za 21. stoljece: nacela i struktura Blissove bibliografske klasifikacije = A classification for the 21st century: principles and structure of the Bliss Bibliographic Classification*, "Vjesnik bibliotekara Hrvatske", 44 (2001), 1/4, p. 38-51).

Essential classification, London, Facet, 2004.

Essential thesaurus construction, London, Facet, 2006 (in corso di pubblicazione in lingua italiana: *Costruire thesauri*, a cura di Piero Cavaleri, traduzione di Laura Ballestra e Luisa Venuti, Editrice Bibliografica, 2008).

Facet analysis as a fundamental theory for structuring subject organization tools, NKOS Workshop of the European Conference on Digital Libraries, Budapest, 17-21 September, 2007, <<http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2007/presentations/BROUGHTON3.PPT>>.

A faceted classification as the basis of a faceted terminology, "Axiomathes" (in corso di pubblicazione). [DOI 10.1007/s10516-007-9027-7]

Henry Evelyn Bliss: the other immortal or a prophet without honour?, "Journal of Librarianship and Information Science", 40 (2008), 1, p. 43-58. [DOI 10.1177/0961000607086620]

Meccano, molecules and knowledge organization: the continuing contribution of S.R. Ranganathan, ISKO UK event "Ranganathan revisited: facets for the future", University College London, 5 November, 2007, <http://www.iskouk.org/presentations/vanda_broughton.pdf>.

The need for a faceted classification as the basis of all methods of information retrieval, "Aslib Proceedings", 58 (2006), 1/2, p. 49-72.

tuto fare, con pochi esempi pratici. I problemi per indicizzatori e classificatori erano ampiamente collegati all'organizzazione fisica delle collezioni, e al modo in cui argomenti complessi potevano essere ricondotti ad un ordine lineare, non solo sugli scaffali di libri ma anche nelle schede cartacee del catalogo, nelle bibliografie cartacee e negli indici di periodici.

Lo sviluppo dell'analisi a faccette fornì un sistema molto lineare per rappresentare l'informazione semantica in un modo logico e prevedibile. Questo aiutò molto sia in fase di indicizzazione sia nel recupero perché, per quanto complicato fosse l'argomento, l'indicizzatore conosceva le regole per sintetizzare le notazioni a partire dagli elementi costitutivi dell'argomento, e poteva prevedere in quale punto della sequenza i documenti dovessero essere collocati.

Ma questa idea andava molto più lontano, perché la struttura che dava ai dati si dimostrò anche molto adatta a un contesto non lineare. Ciò apparve evidente quando negli anni Ottanta comparvero i database bibliografici e i primi cataloghi elettronici (che sono solo un tipo specifico di database bibliografici). Fu probabilmente allora che si cominciò a ragionare sulla struttura dei dati, semplicemente perché doveva essere compresa dalle macchine, e quindi più coerente e regolare. Oggi, con l'avvento di Internet e nel contesto delle risorse elettroniche, i problemi di dare un ordinamento lineare hanno perso importanza, perché non c'è più nessuna collezione fisica da ordinare, anche se in questo contesto rimane il bisogno di pensare come presentare le risorse e gli strumenti di organizzazione della conoscenza che sono ancora rilevanti.

Ritengo che un'altra peculiarità della situazione attuale, importante da sottolineare, sia quella che le informazioni spesso non si presentano

nei comodi e semplici formati di trent'anni fa, cioè sotto forma di libro o articolo di periodico. Oggi gli oggetti a cui noi dobbiamo "incollare" le etichette dei nostri indici possono essere complessi e diffusi e, naturalmente, possono anche cambiare nel tempo, vista la massima facilità con cui le risorse online possono essere corrette e aggiornate.

Come risultato di ciò, le risorse tendono ad essere meno stabili e infatti versioni differenti possono scomparire completamente, in modi che mai sarebbero stati possibili per le risorse cartacee.

Oltre a questi cambiamenti della forma dell'informazione, le discipline sono più soggette a reinterpretazione e ridefinizione, nuove discipline sorgono non solo attraverso nuove ricerche e per la creazione di nuova conoscenza, ma anche per la crescita di studi "cross-disciplinari" e interdisciplinari e per l'adozione di metodologie mutuata da altre discipline.

Questa natura diffusa e indefinita dell'informazione significa che gli attuali strumenti di indicizzazione hanno bisogno di essere particolarmente flessibili e reattivi se vogliono essere efficaci. È molto gratificante pensare che la teoria dell'indicizzazione, in particolar modo quella della teoria dell'analisi a faccette, si dimostra ancora solida in un contesto come questo ed è

adattabile ad una situazione che pone esigenze tanto alte.

Questo avviene perché l'enfasi è sulla *metodologia* dell'analisi, sull'identificazione di ruoli e categorie, sulla determinazione delle relazioni tra termini e concetti, non sulla creazione di strutture fisse come nel caso di classificazioni enumerative.

Questo tipo di analisi precisa e rigorosa è di sicuro essenziale per un trattamento automatico dei dati e questo è senza dubbio ciò che la rende così attraente per professionisti esterni alle discipline LIS.

Penso sia importante ricordare che i principi dell'indicizzazione e della costruzione dei vocabolari rimangono validi. Sebbene ci si trovi in un contesto molto differente, la teoria fondamentale sviluppata a metà del XIX secolo continua a fornire solide basi per la gestione dell'informazione.

Note

¹ JACK MILLS – VANDA BROUGHTON, *Bliss Bibliographic Classification. Introduction and auxiliary schedules*, Butterworth, 1977, 2. ed. (ndr).

² JEAN AITCHISON – ALAN GOMERSALL – RALPH IRELAND, *Thesurofacet: a thesaurus & faceted classification for engineering & related subjects*, edited by Jean Aitchison [and] Patricia Day, Leicester, English Electric Co. Ltd., 1969 (ndr).

Abstract

Vanda Broughton, editor with Jack Mills of BC2, describes in this interview the story of Bliss Classification revision after 1977 and recent evolutions of BC2 scheme and indexing tools in web context (thesauri, classifications, ontologies, taxonomies...). Thesaurus and classification schemes are regarded as similar kinds of tool, because facet analysis principles could be used to construct both types of vocabulary.

Broughton suggests that many computer scientists are now far more aware of the theoretical issues related to good information architecture and librarians should speak to them in their own language, maintaining a bridge between intellectual, practical and technical aspects of cataloguing and indexing, and ensuring that all these aspects are kept in balance. Even in digital and hybrid libraries the principles of indexing and of constructing vocabularies remain true, and although the environment is very different, the fundamental theory that was developed in the mid twentieth century continues to provide us with a very sound basis for managing information.