

## Rilevanza nel documento e rilevanza per l'utente: l'ordine di citazione dei temi

Claudio Gnoli

### Preprint del contributo al volume

*Viaggi a bordo di una parola: scritti sull'indicizzazione semantica in onore di Alberto Cheti*,  
a cura di Anna Lucarelli, Alberto Petrucciani, Elisabetta Viti, AIB, Roma 2019, p. 105-117

### English abstract

*Relevance in the document and relevance for the user: the citation order of themes.* Alberto Cheti and his Research Group on Subject Indexing (GRIS) have been inspirational sources in introducing the notion of *theme* in theory of subject indexing. A document has a base theme and possibly some particular themes. Within a subject heading or a classification number, the base theme should be cited first. This has important consequences for its ordering in browsable lists of subjects. Some users may be interested in particular themes of documents expressed after their base theme. Search and browse interfaces should suggest such resources by separate "see also" references, that can be produced through a double query method. A proposed scheme of dimensions, including reality, phenomena, perspectives, documents, collections, information needs and people, may be of help to the analysis.

### 1: Introduzione

Fin dalle prime occasioni in cui ho avuto la fortuna di incontrarlo, di Alberto Cheti mi ha colpito la profondità originale di pensiero. In mezzo alle tante voci che discutevano i temi d'attualità nella biblioteconomia del momento, i suoi interventi hanno sempre suggerito, con modi inusualmente modesti e garbati, prospettive più sostanziali secondo le quali i diversi aspetti dell'indicizzazione possono venire considerati.

Questa impressione si è andata confermando anche man mano che ho potuto confrontarla con le correnti principali della discussione internazionale nel dominio più ampio dell'organizzazione della conoscenza (*knowledge organization, KO*), entro il quale i principî bibliotecari dell'indicizzazione per soggetto tendono oggi ad essere ricompresi. Mi sono cioè reso conto che l'esperienza sviluppatasi in Italia attorno al Gruppo di ricerca sull'indicizzazione per soggetto (GRIS), del quale Alberto è ispiratore essenziale, aveva raggiunto livelli di analisi più sofisticati rispetto alla vulgata diffusa nei testi di lingua inglese, specialmente quella influenzata dalla potenza organizzativa delle istituzioni statunitensi e riflessa nello sviluppo dei Library of Congress Subject Headings (LCSH) e della Classificazione decimale Dewey (CDD) e nell'elaborazione degli standard dell'IFLA e del W3C. La scuola italiana – e toscana in particolare – sarebbe in grado di apportare contributi nuovi ed importanti alla teoria internazionale, se questi riuscissero a superare certe barriere linguistiche e culturali che invece spesso portano a un'eccessiva semplificazione dei principî concettuali in gioco.

Un esempio importante è la nozione di *tema* approfondita nell'ambito del GRIS, che Bultrini e Cheti (2008) hanno tentato di esportare nella discussione sullo sviluppo dei modelli concettuali di FRBR, ma sembra per il momento essere stata recepita dall'IFLA più a un livello terminologico, con la distinzione fra *thema* e *nomen* (Žumer 2017, par. 4), che nella rappresentazione analitica della struttura dei possibili soggetti di un documento.

Dedicherò dunque questo mio omaggio ad Alberto alle relazioni fra tale nozione e i principî correnti della classificazione e della soggettazione, insieme all'auspicio che l'insieme del suo lavoro possa essere adeguatamente riconosciuto ed applicato ai livelli assai generali che gli competono. A motivare questo augurio concorrono tanto il valore intrinseco quanto l'affezionato ricordo dei nostri

scambi, fattisi più intensi con gli incontri sul Nuovo Soggettario e l'ILC, la partecipazione alla vita del capitolo italiano dell'ISKO, la collaborazione a pubblicazioni estere e i dialoghi pieni di umanità e amicizia insieme alle rispettive famiglie fra le colline fiorentine.

## 2: Tema di base e temi particolari

In linguistica testuale, come Alberto ci ha insegnato (Cheti 1996), il *tema* di una frase o di un intero testo è ciò di cui questi intendono parlare, che come tale introduce il contesto del discorso da sviluppare. Nella frase "Per quanto riguarda il programma di domani, lo decideremo anche in base al tempo atmosferico", il programma di domani è il tema; il resto della frase costituisce invece il *rema*, ovvero ciò che si dice a proposito del tema. In analisi logica tema e rema corrispondono in genere a soggetto e predicato: in "le tigri sono pericolose" le tigri sono il tema, e il loro essere pericolose è il rema. Non a caso, nella maggior parte delle lingue il soggetto-tema viene posto all'inizio di una frase, e il rema lo segue (se occorre invece tematizzare un complemento si ricorre ad apposite costruzioni invertite come in "il pacchetto, l'ho portato con me"). Questa considerazione, come vedremo, assume un importante valore cognitivo anche nelle applicazioni all'indicizzazione per soggetto.

Considerando interi documenti, come facciamo di solito nell'indicizzazione, il tema corrisponde spesso a quanto viene espresso dalle sue parti iniziali e dalle sue sintesi, come il titolo, il sottotitolo, il sommario e la quarta di copertina di un libro o il trailer di un film. Il tema ci dice di che cosa si parlerà; se poi vogliamo sapere che cosa specificamente se ne dirà, dovremo leggere il libro o guardare il film. La teoria dell'indicizzazione in genere assume che sia utile riferirsi a un documento indicandone l'*aboutness*, ossia appunto il tema, sebbene qualche autore abbia suggerito che utenti molto specializzati, che già conoscono bene la natura dei temi in questione, sarebbero meglio serviti da un'indicizzazione che comprenda anche i remi (Weinberg 1988).

La nozione di tema è stata introdotta dal GRIS nella definizione del soggetto di un documento: "In linea generale e a prescindere da contesti particolari, soggetto può essere considerato ogni concetto o combinazione di concetti che rappresentino un tema all'interno di un documento. Tuttavia, per evitare incongruenze nei risultati dell'analisi, è opportuno che il soggetto sia sempre principalmente identificato con il *tema di base* del documento" (GRIS 1996, par. 1.2).

Infatti Hjørland (2017, par. 3.6), seguendo uno standard ISO e altre fonti, considera *theme* come uno tra i sostanziali sinonimi di *subject*, insieme a *aboutness*, *topic*, *isness* e *ofness*. Tuttavia la definizione del GRIS introduce immediatamente una distinzione fra tema di base e temi particolari, che nel loro insieme permettono di considerare il soggetto in un modo più analitico: "Il tema di base è quell'oggetto unitario di conoscenza al quale sono riferibili i singoli temi particolari discussi nel documento e al quale sono correlate nel testo tutte le informazioni fornite intenzionalmente dall'autore, essendo stata proprio la volontà di comunicare nozioni dirette e specifiche su quell'oggetto di conoscenza il motivo fondamentale della produzione intellettuale dell'intero documento" (GRIS 1996, par. 1.2).

Il soggetto di un documento potrà dunque comporsi di un tema di base e di uno o più temi particolari. L'analisi concettuale del documento dovrebbe condurre all'identificazione del suo tema di base ed eventualmente di temi particolari che l'impianto generale della trattazione riferisce ad esso. L'indicizzazione per soggetto del documento dovrà basarsi innanzitutto sul tema di base, come raccomanda la *Guida* GRIS; tuttavia è significativo il riconoscimento che il soggetto possa essere un'entità complessa, nella quale coesistono componenti diversi. Anche questi ultimi infatti possono svolgere un ruolo utile nell'organizzazione e nel reperimento dei documenti.

Considerando per esempio l'intervento dello stesso Alberto Cheti al seminario su *L'indicizzazione* svoltosi nel 2002 a Modena (Cheti 2003), si può riscontrare che esso tocca diversi temi fra loro collegati: (1) lo sviluppo – allora iniziale – del Nuovo Soggettario, (2) la traduzione italiana della CDD, (3) l'applicazione del PRECIS alla lingua italiana e (4) la definizione di soggetto secondo il GRIS. La struttura dell'articolo suggerisce appunto il fatto che fra questi temi

esistano connessioni in quanto, pur trattandosi di esperienze distinte che hanno prodotto strumenti di indicizzazione diversi, esse si sono reciprocamente influenzate attraverso l'emersione di principi sostanzialmente comuni di KO. Un'analisi concettuale dell'articolo potrebbe condurre a scegliere *Nuovo Soggettario – sviluppo* come il suo tema di base, registrando comunque gli altri come temi particolari che hanno in vario modo un'influenza sul tema di base.

Nell'indicizzazione per soggetto verbale, autori come Revelli (1986, citato in Cheti 2003) hanno raccomandato che la formulazione finale del soggetto consista in una sola "stringa unica coestesa", il che implicherebbe una selezione dei temi particolari ed eventualmente un loro riferimento al tema di base mediante connessioni sintattiche all'interno della stessa stringa. Tuttavia le scelte possono essere diverse a seconda degli specifici contesti e iniziative di indicizzazione. Nella traduzione del soggetto in simboli di una classificazione, esso non viene generalmente discusso in termini di tema di base e temi particolari, ma come vedremo i rapporti fra i concetti componenti il soggetto possono essere analoghi; perciò la nozione di tema è stata introdotta, ispirandosi anche agli insegnamenti di Alberto, nella struttura dell'Integrative Levels Classification (ILC, Gnoli 2018).

L'uso tradizionale delle classificazioni per collocare libri sugli scaffali prevede che questi siano assegnati ad una sola classe, rappresentativa del tema di base, tuttalpiù specificata da suddivisioni comuni o da faccette. Se, per considerare un caso comune nella mia esperienza, un volume tratta di dinamica dei fluidi e delle sue applicazioni in ingegneria idraulica, l'impostazione disciplinare della CDD richiede che il catalogatore decida se sia prevalente il tema fisico oppure quello ingegneristico, e scelga di conseguenza fra la classe 532 e la classe 627. D'altronde, una singola copia del volume non potrebbe stare contemporaneamente su entrambi gli scaffali. Nelle applicazioni ai cataloghi e ai repertori, tuttavia, la situazione è diversa: sul piano tecnico nulla impedisce di assegnare ad una stessa risorsa più di un numero di classe, permettendo così di recuperarla sia attraverso una ricerca di documenti sulla meccanica dei fluidi che attraverso una sull'ingegneria idraulica. Nonostante una certa resistenza nelle istruzioni ufficiali a considerare queste interpretazioni, non è raro osservare applicazioni che associano diverse classi CDD ad uno stesso documento. Questa scelta è adottata sistematicamente nel Basel Register of Thesauri, Ontologies and Classifications (BARTOC), il repertorio di sistemi per l'organizzazione della conoscenza (KOS) che assegna a ciascuno di essi, quale elemento della sua indicizzazione, anche diverse classi CDD. Così ad esempio la ProdCom List, una classificazione delle Comunità europee per le statistiche produttive di miniere, cave e susseguenti lavorazioni, è indicizzata sia con 338 "produzione" che con 622 "arte mineraria e attività connesse", e dunque trovabile sia attraverso le gerarchie dell'economia che attraverso quelle dell'ingegneria.

In queste applicazioni i diversi temi associati a una risorsa sono semplicemente giustapposti l'uno all'altro, senza che sia previsto alcun tipo di relazione sintattica fra di loro. Tuttavia alcuni sistemi di classificazione permettono di combinare classi diverse attraverso specifiche relazioni, che Ranganathan ha chiamato *relazioni di fase* (Bianchini 2017). Nella Classificazione decimale universale (CDU) il simbolo per queste relazioni sono i due punti; così 338:622 significa "produzione in relazione all'arte mineraria". Simile è la funzione del punto e virgola nella classificazione elaborata da Ingetraut Dahlberg per la documentazione sull'organizzazione della conoscenza, tuttora applicata nel repertorio online *Knowledge organization literature* <<http://www.isko.org/lit.html>>. Per esempio, l'articolo "A grammatical approach to subject classification in museums" (Szostak 2017) è classificato 225;878, dove 225 significa "faccette" e 878 "classificazione e indicizzazione di oggetti museali".

Oltre alla relazione di fase generale (*0a* nella Classificazione Colon), Ranganathan prevede la possibilità di specificare relazioni di orientamento di una fase verso un'altra (*bias*, *0b*), di confronto fra due fasi (*0c*), di differenza fra due fasi (*0d*), di influenza di una fase su un'altra (*0g*). Queste relazioni ci appaiono analoghe a quelle che possono stabilirsi fra i diversi temi di un soggetto composto. Ad esempio, l'ultima potrebbe esprimere l'influenza del tema particolare "traduzione in italiano della CDD" sul tema di base "sviluppo del Nuovo Soggettario".

### 3: L'ordine di citazione dei temi e i suoi effetti

Nel momento in cui una stringa di soggetto o una classe siano composte di più temi o fasi, si pone automaticamente la questione dell'ordine nel quale essi debbano essere citati. Per la relazione di fase generale, Ranganathan prescrive che si citi per prima la classe che compare per prima nelle tavole (dunque quella di minor concretezza o minor artificialità, in base ai principi della Colon), criterio che ha più che altro lo scopo di assicurare una coerenza fra soggetti affini assegnati a documenti diversi. Tuttavia, considerando i principi del GRIS, il primo elemento di un soggetto dovrebbe corrispondere sempre al tema di base: esso cioè dovrebbe dipendere non dalla struttura dello schema utilizzato, né dal potenziale interesse per gli utenti, bensì da come i temi sono trattati all'interno del documento stesso, "essendo stata proprio la volontà di comunicare nozioni dirette e specifiche su quell'oggetto di conoscenza il motivo fondamentale della produzione intellettuale dell'intero documento". Pertanto, la scelta del primo tema da citare dovrebbe essere operata in base all'esame di ciascun documento. L'ordine 225;878 esprime il fatto che nell'articolo di Szostak l'analisi a faccette sia tema di base rispetto all'indicizzazione museale, che assume il ruolo di tema particolare. Se invece il tema di base si rivelasse l'indicizzazione museale, allora la classe composta andrebbe scritta 878;225.

Queste norme potrebbero sembrare ormai superflue a chi è avvezzo al funzionamento dei sistemi informatici di recupero dell'informazione, poiché è noto che l'elaboratore, applicando la logica booleana, sarà in grado di recuperare la scheda bibliografica tanto cercando 225 quanto cercando 878. Questo però è vero soltanto per una delle funzioni fondamentali della ricerca bibliografica (Žumer 2017, par. 3): *find*, ossia "radunare informazioni su una o più risorse di proprio interesse cercando in base a qualsiasi criterio pertinente"; lo è meno per le altre funzioni *identify*, *select* e soprattutto *explore* (tralasciamo qui la funzione pratica *obtain*). Infatti, *identificare* una risorsa significa nella definizione dell'IFLA "comprenderne la natura e distinguere fra risorse simili", ad esempio fra risorse in cui il tema di base è l'indicizzazione museale in relazione al tema particolare delle faccette e risorse in cui vale l'inverso. Altrettanto vale per *select* ossia "stabilire l'idoneità delle risorse trovate ed essere messi in grado di accettare o scartare singole risorse": se all'utente interessa l'indicizzazione museale, saranno assai più idonee le risorse che la trattano come tema di base rispetto a quelle che la trattano solo come tema particolare.

Quanto poi a *explore* ossia "scoprire delle risorse utilizzando le relazioni fra loro e collocare così le risorse entro un contesto", sappiamo quanto sia cruciale a questo scopo una disposizione ordinata di tutti i soggetti disponibili, oppure dei soli soggetti recuperati con una prima ricerca, affinché l'utente possa appunto esplorare le potenziali fonti collocandole entro il contesto più ampio del sistema di organizzazione della conoscenza. Questa funzione viene consentita dai criteri di ordinamento che assicurano coerenza al KOS, ossia l'ordinamento alfabetico nei sistemi verbali e quello sistematico nei sistemi classificati. È evidente come la scelta dell'elemento citato per primo, che dovrebbe corrispondere al tema di base, risulti determinante per la posizione che uno specifico soggetto assumerà tanto in un ordinamento alfabetico quanto in uno sistematico. Mentre 225;878 verrà elencato insieme ad altre risorse che trattano di faccette, 878;225 verrà elencato con risorse che trattano di musei. I semplici processi di raggruppamento e giustapposizione operati automaticamente dall'elaboratore avranno effetti cognitivamente assai diversi nei due casi.

Appare qui chiaro quanto sia importante distinguere fra la rilevanza di un tema rispetto agli altri temi all'interno di un documento (tema di base vs. temi particolari) e la rilevanza di un tema rispetto agli altri per uno specifico utente. Nonostante l'obiettivo ultimo dei servizi bibliografici sia quello di giovare ai loro utenti, non è naturalmente possibile conoscere in anticipo la rilevanza per ogni utente, quella che associa "ad ogni lettore il suo libro", dal momento che essa muta ad ogni ricerca. Di conseguenza, occorre che il KOS venga applicato ai documenti in modo imparziale, evidenziando la rilevanza di ciascun tema solo rispetto al documento. La seconda fase della discriminazione fra le risorse potrà infatti essere compiuta solo al momento della ricerca, sfruttando le caratteristiche del KOS e quelle della sua interfaccia.

Qualche deroga a questo principio viene suggerita dal GRIS laddove consiglia di esprimere in un soggetto, tra gli svariati temi particolari di un documento, soltanto quelli di presunto interesse per gli utenti: ma in questo modo le due funzioni informative vengono mescolate. In ogni caso, la scelta del tema di base dipende rigorosamente dal documento e non dagli utenti, come è stato già evidenziato sopra. Più ragionevole ci appare il tener conto dell'interesse degli utenti al momento di assegnare al documento una collocazione. Qui è infatti possibile sdoppiare l'informazione fra quella sul soggetto del documento in sé, da registrare nei campi del catalogo dedicati alla classificazione o al soggetto verbale, e quella sulla sua collocazione fisica, che può anch'essa essere basata su una classificazione. Ad esempio un libro in cui il tema di base sia la meccanica dei fluidi, con esempi di applicazioni all'ingegneria idraulica (tema particolare), si potrà classificare come 532:626, ma collocare in 626 a beneficio degli utenti della propria particolare biblioteca, se si presume che questi siano più interessati agli aspetti ingegneristici che a quelli fisici di quel documento.

#### 4: Le dimensioni nell'organizzazione della conoscenza

A proposito delle interazioni fra soggetti e utenti può risultare utile la schematizzazione da me proposta (Gnoli 2016) delle diverse dimensioni implicate nei processi di organizzazione della conoscenza. Lo schema sviluppa una nota di Brian Vickery circa i successivi passaggi "dal mondo fino al classificatore", e propone la distinzione di una serie di dimensioni contrassegnate da lettere greche:

- $\alpha$  realtà in sé
- $\beta$  fenomeni
- $\gamma$  prospettive
- $\delta$  documenti
- $\epsilon$  collezioni
- $\zeta$  bisogni informativi
- $\eta$  persone

Non essendoci concesso attingere alla realtà in sé in modo diretto, ma solo attraverso i nostri sensi e gli altri strumenti impiegati dalla ricerca, le unità di riferimento della conoscenza umana devono essere i fenomeni ( $\beta$ ), ossia le manifestazioni della realtà come la conosciamo attualmente, risultanti dall'interazione fra i concetti già formati e i percetti che man mano ci provengono dalle nuove esperienze. Per esempio, il fenomeno "tigri" approssima una qualche realtà ultima delle specie delle tigri così come finora siamo riusciti a comprenderle e classificarle.

Le tigri possono poi essere studiate secondo diverse discipline, approcci, teorie, metodi e via dicendo, che costituiscono nel loro insieme la dimensione delle prospettive ( $\gamma$ ). Di tigri può cioè occuparsi la medicina veterinaria così come l'arte ceramica cinese o la narrazione di una tigre immaginaria nel geniale *Calvin & Hobbes*. Le classificazioni organizzate per discipline, come la CDD o la Colon, elencano dunque innanzitutto delle classi di prospettive, mentre i fenomeni da esse studiati si ritrovano in punti diversi delle tavole a seconda del contesto disciplinare (tigri in veterinaria, tigri in letteratura).

La conoscenza elaborata da ciascuna prospettiva è poi espressa in documenti ( $\delta$ ) che possono assumere i più svariati formati, lingue, tecniche di rappresentazione e via dicendo: documentari video sulle tigri, tavole di tigri disegnate, in forma analogica oppure digitale... Tema di base e temi particolari, come abbiamo visto, sono caratteristiche di uno specifico documento, come dovrebbe essere rappresentato in una bibliografia o un repertorio. Così "tigri nell'arte ceramica cinese" sarà una combinazione di fenomeno e prospettiva che è tema di un certo documento.

I documenti sono raccolti e organizzati in collezioni ( $\epsilon$ ) generali oppure specializzate, quali archivi, biblioteche, orti botanici, musei e gallerie. A questo livello, si è detto, il catalogo sceglie di evidenziare alcuni temi di presunto interesse per l'utenza della collezione, ma nel caso delle

collezioni pubbliche dovrebbe mantenere una certa neutralità, non potendo prevedere in anticipo l'identità di ciascun utente. Le collezioni sono infatti fruite da una molteplicità di persone ( $\eta$ ), ossia i loro utenti, con diverse caratteristiche sociali, familiari e personali.

Alla discussione seguita alla presentazione di questo schema (Ridi 2015) ad un incontro di ISKO Italia, Alberto Cheti ha partecipato con la consueta profondità, suggerendo l'inserimento di una dimensione aggiuntiva che renda conto dei bisogni informativi degli utenti. Ho accolto volentieri l'idea contrassegnandola come dimensione  $\zeta$  (che Riccardo Ridi, coautore di quell'intervento, avrebbe allora preferito chiamare dei "comportamenti informativi", includendovi "i bisogni ma anche i vincoli, i condizionamenti, le abitudini, le azioni di ricerca, diffusione, valutazione, fruizione, organizzazione, conservazione, ecc.", come mi espresse in un messaggio di posta elettronica).

La dimensione chetiana  $\zeta$  appare specialmente rilevante nel discutere la rilevanza dei temi per gli utenti: infatti uno stesso utente  $\eta$  può avvicinarsi ad una collezione con bisogni informativi diversi a seconda delle motivazioni che ve lo spingono in quel momento: portare avanti la propria tesi di laurea o trovare una lettura per le vacanze, aiutare un amico o perfino dar sfogo al proprio umore del momento, diverso dall'umore del giorno precedente. Assumere che un tema sia sempre rilevante o sempre irrilevante per un certo utente significherebbe in fondo categorizzare le persone imponendo loro griglie sociologiche troppo rigide: in una società aperta, una biblioteca "per ragazzi" non dovrà per questo contenere solo libri di scuola e fumetti, poiché esistono ragazzi curiosi di sviluppare in proprio altri argomenti; ciò che una collezione deve puntare a soddisfare sono innanzitutto certi bisogni informativi, indipendentemente da chi sia ad averli.

## 5: Temi e recupero delle informazioni: il metodo della doppia query

Se dunque la rilevanza dei temi nel documento dev'essere espressa nell'ordine di citazione interno a un soggetto, ponendo per primo il tema di base seguito eventualmente dai temi particolari, la rilevanza dei temi per l'utente e i suoi bisogni informativi del momento deve invece essere gestita dalle interfacce degli indici. Questa separazione di funzioni permetterà di ottimizzare la precisione in entrambe le dimensioni,  $\delta$  e  $\zeta$ .

Consideriamo il libro *L'ambiente: conoscerlo e proteggerlo: percorsi di chimica ambientale* (Campanella - Conti 2010). La sua scheda nell'opac nazionale SBN <<http://opac.sbn.it/bid/UBO3825804>> gli assegna due soggetti, corrispondenti a due temi: *Ambiente naturale – protezione* e *Chimica ambientale*. La classe della 21esima edizione della CDD che gli è stata assegnata è 628.50154 "Tecnologia della tutela ambientale. Tecnologia del controllo dell'inquinamento. Chimica", appartenente alla gerarchia dell'ingegneria; è evidente però che si tratta di un'opera interdisciplinare e la CDD, che non dispone di mezzi per esprimere le relazioni di fase, riesce almeno ad utilizzare la suddivisione comune -0154 "Principi chimici" per introdurre il tema particolare della chimica (purtroppo questo dispositivo esiste solo per le relazioni con sottoclassi delle scienze).

La CDU potrebbe meglio esprimere la giustapposizione dei temi con 628.5:54 "misure contro gli inquinanti industriali e altri, in relazione con la chimica". In essa un altro documento sugli stessi temi, ma nel quale la chimica abbia il ruolo di tema di base, si esprimerebbe con 54:628.5. Questi e altri soggetti potrebbero essere presenti in un elenco di documenti ordinati per classe:

54	chimica
54:628.5	chimica: misure contro gli inquinanti industriali
547	chimica organica
547:628.5	chimica organica: misure contro gli inquinanti industriali
...	

62	ingegneria
628	ingegneria della salute pubblica
628.5	misure contro gli inquinanti industriali
628.5:54	misure contro gli inquinanti industriali: chimica
628.5:57	misure contro gli inquinanti industriali: biologia
...	

Un utente motivato alla ricerca di informazioni sugli inquinanti industriali, guidato dalle gerarchie della CDU, arriverà presumibilmente alla zona della classe 628.5 e non alla zona della classe 54. Attivando un collegamento ipertestuale da 628.5 otterrà un elenco di documenti che hanno per tema di base le misure contro gli inquinanti industriali, da solo o combinato con altri temi (in relazione alla chimica, alla biologia, ecc.). Fra questi vi sarà, appropriatamente, il libro di Campanella e Conti. Se questo fosse stato indicizzato con un diverso ordine di citazione dei temi (54:628.5), non comparirebbe in tale elenco.

D'altra parte, il nostro utente potrà essere interessato anche a documenti che trattino delle misure contro gli inquinanti industriali come tema particolare, anche se essi saranno presumibilmente meno rilevanti per lui di quelli in cui tali misure sono il tema di base. Una buona interfaccia dovrebbe allora suggerire all'utente, dopo avergli presentato i primi risultati, un rinvio del tipo "vedi anche" agli altri documenti potenzialmente rilevanti, nei quali 628.5 compare come tema particolare e dunque in posizione diversa dalla prima.

Abbiamo suggerito (Gnoli - Cheti 2013) che questa funzione di suggerimento, estremamente rara nelle interfacce a nostra conoscenza, richiede la programmazione dell'interfaccia con un metodo costituito da una doppia query. In altre parole, la selezione della classe 628.5 dovrebbe lanciare nella base-dati due successive ricerche: la prima per le registrazioni in cui 628.5 compaia all'inizio della stringa del numero di classificazione, e la seconda per quelle in cui 628.5 compaia in posizione diversa dall'inizio. Questo permetterà di visualizzare i risultati in due sezioni, cominciando da quelli nei quali la classe selezionata è proprio il tema di base, e che quindi sono potenzialmente più rilevanti. La maggior parte delle interfacce invece applica una sola query, che restituisce o solo il primo gruppo di risultati (qualora il troncamento del numero di classe sia impostato solo a destra: 628.5\*) oppure l'insieme di tutti i risultati mescolati (qualora il troncamento sia impostato anche a sinistra: \*628.5\*). In questo secondo caso però l'ordinamento sistematico – che è comunque quello più raccomandabile – presenterebbe per prime le classi come 54:628.5, che precedono 628... nell'ordine dello schema, ma sono meno appropriate ai fini della rilevanza per l'utente.

Il metodo della doppia query applicata a classi composte di diverse fasi è stato implementato nell'interfaccia della *Knowledge organization literature* citata sopra, che può essere sperimentata per averne una dimostrazione pratica.

## 6: Conclusione e prospettive

La rapida rassegna che abbiamo proposto dovrebbe mostrare come la nozione di tema, introdotta dal GRIS nella teoria dell'indicizzazione per soggetto, sia feconda di implicazioni. Sovrapponendosi solo parzialmente alle nozioni di *subject*, *aboutness* ecc. essa permette un'analisi più adeguata del soggetto e delle sue componenti, con la distinzione fra tema e rema e quella fra tema di base e temi particolari. Può essere applicata non soltanto all'indicizzazione per soggetto verbale ma anche alla classificazione, dove risulta affine alla nozione di fase. In entrambi i casi ha importanti implicazioni per l'ordine di citazione nei soggetti complessi, di cui dovrebbero tenere conto sia l'attività di indicizzazione sia la progettazione delle interfacce di ricerca ed esplorazione dei cataloghi.

L'individuazione dei temi di un documento e delle relazioni che fra loro intercorrono permette di rappresentare la "macrostruttura" del testo nella "microstruttura" di una voce di

soggetto, composta di elementi concettuali collegati fra loro da relazioni. Questo modo di descrivere i soggetti lascia intravedere anche una generalizzazione dei tipi di relazioni fra concetti, relazioni che siamo finora abituati a distinguere in faccette, interne a un tema, e relazioni di fase, fra temi: in quello che Alberto ha descritto come "il metodo dell'analisi categoriale, in un'accezione più ampia rispetto alle classificazioni a faccette" (Cheti 2003, cfr. Cheti 1990), tra il significato semantico di una faccetta d'agente e quello di una relazione di fase di influenza non sembra esserci una differenza sostanziale, sicché una sintassi *a faccette libere* come quella raccomandata da Austin (1976) permetterebbe di trattarle allo stesso modo, distinguendo soltanto diversi gradi di annidamento nella struttura sintattica di un soggetto. In questo quadro, la nozione chetiana di tema si propone come un elemento assai generale, potenzialmente in grado di sussumere numerose funzioni in una rinnovata teoria dell'organizzazione della conoscenza.

## Riferimenti bibliografici

Austin Derek, *The CRG research into a freely faceted scheme*. In: *Classification in the 1970s: a second look*, edited by Arthur Maltby. London: Bingley, 1976, p. 158-194.

Bianchini Carlo, 2017, Il trattamento delle relazioni nella classificazione Colon, «JLIS.it», 8 (2017), no. 2, p. 52-68.

Campanella Luigi; Conti Marcelo Enrique, *L'ambiente: conoscerlo e proteggerlo: percorsi di chimica ambientale*. Roma: Carocci Faber, 2010.

Cheti Alberto, *Le categorie nell'indicizzazione*, «Biblioteche oggi», 8 (1990), n. 1, p. 29-49.

Cheti Alberto, *Manuale ipertestuale di analisi concettuale*, realizzazione di Serena Spinelli. Università di Bologna, <[https://biblioteche.unibo.it/manuals/html\\_1/HOME.HTML](https://biblioteche.unibo.it/manuals/html_1/HOME.HTML)>, 1996 (in particolare <[html\\_1/tema.htm](html_1/tema.htm)>).

Cheti Alberto, *Il punto sul GRIS e gli sviluppi attuali*, «Bibliotime», 6 (2003), n. 1, <<http://www.aib.it/aib/sezioni/emr/bibtime/num-vi-1/cheti.htm>>.

Gnoli Claudio; Cheti Alberto, *Sorting documents by base theme with synthetic classification: the double query method*. In: *Classification & visualization: interfaces to knowledge: proceedings of the international UDC Seminar, 24-25 October 2013, The Hague*, eds. Aida Slavic, Almila Akdag Salah, Sylvie Davies. Würzburg: Ergon, 2013, p. 225-232.

Gnoli Claudio, *Classifying phenomena, part 1: Dimensions*, «Knowledge organization», 43 (2016), n. 6, p. 403-415.

Gnoli Claudio, *Classifying phenomena, part 4: Themes and rhemes*, «Knowledge organization», 45 (2018), n. 1, p. 43-53.

GRIS: Gruppo di ricerca sull'indicizzazione per soggetto, *Guida all'indicizzazione per soggetto*. Roma: AIB, 1996. Versione WWW a cura di Eugenio Gatto, <<http://www.aib.it/aib/gris/gris.htm>>.

Hjørland Birger, *Subject (of documents)*. In: *ISKO encyclopedia of knowledge organization*, ed. Birger Hjørland, <<http://www.isko.org/cyclo/subject>>, 2016. Anche in «Knowledge organization», 44 (2017), n. 1, p. 55-64.



Revelli Carlo, *Il catalogo per soggetti e le aspettative dei bibliotecari nei confronti dell'automazione*. In: *Il recupero dell'informazione: atti del Convegno-esposizione bibliografica Indicizzazione per soggetto e automazione, Trieste, 21-22 ottobre 1985*, a cura di Adriano Dugulin, Antonia Ida Fontana, Annamaria Zecchia. Milano: editrice Bibliografica, 1986, p. 27-53.

Ridi Riccardo, *Livelli di irrealità: oggettività e soggettività nell'organizzazione della conoscenza*. «Bibliotime», 18, n. 2, <<http://www.aib.it/aib/sezioni/emr/bibttime/num-xviii-2/ridi.htm>>.

Szostak Rick, *A grammatical approach to subject classification in museums*, «Knowledge organization», 43 (2017), n. 7, p. 494-505.

Weinberg Bella Hass, *Why indexing fails the researcher*, «The indexer», 16 (1988), n. 1, p. 3-6.

Žumer Maja, *IFLA Library Reference Model (LRM): harmonisation of the FRBR family*. In: *ISKO encyclopedia of knowledge organization*, ed. Birger Hjørland, <<http://www.isko.org/cyclo/lrm>>, 2017.