

L'arte dimenticata dell'ordinamento

Perché non tutto deve andare dalla A alla Z

Claudio Gnoli

Università di Pavia. Biblioteca della scienza e della tecnica

gnoli@aib.it
mate.unipv.it/gnoli

(VI Summit di architettura dell'informazione, Milano, 10-11 maggio 2012)

Abstract. Molte tecniche di organizzazione della conoscenza sono state sviluppate nel corso del Novecento nell'ambito della biblioteconomia e della documentazione: oggi l'architettura dell'informazione le eredita e le applica a nuove funzioni. Fra i principi della classificazione bibliografica, ce n'è però uno che appare dimenticato e rimasto estraneo alla cultura informatica: l'ordinamento sistematico delle voci, alternativo e complementare a quello alfabetico. Sul Web, le voci di ciascun ramo di una tassonomia sono quasi sempre presentate in ordine alfabetico, sebbene un ordine sistematico possa offrire vantaggi. La gestione di tali ordinamenti è possibile mediante sistemi di notazione, gestibili con campi dedicati nelle basi-dati, anche invisibili agli utenti. Vengono presentate e discusse alcune situazioni in cui l'integrazione di una notazione può arricchire le possibilità di scorrimento e visualizzazione di liste di voci, specialmente quando lunghe e strutturate.

Keywords. Organizzazione della conoscenza; Classificazione; Notazione; Ordine alfabetico

1 L'organizzazione della conoscenza

L'uomo si pone fin da tempi assai antichi il problema di come organizzare la massa di conoscenze in suo possesso in schemi logici e facilmente fruibili. Vi si dedicarono infatti già Aristotele, Bacone e i curatori dell'*Encyclopédie* illuminista.

Nel corso del Novecento, le tecniche di organizzazione dei contenuti del sapere si sono sviluppate soprattutto nelle biblioteche e nei centri di documentazione, dove è necessario catalogare e disporre centinaia di migliaia di documenti in modi funzionali, possibilmente raggruppando insieme testi affini.

Oggi l'insieme di queste tecniche, dei principi che ne stanno alla base e degli schemi che utilizzano viene chiamato *organizzazione della conoscenza* – in inglese *knowledge organization (KO)* da non confondere con i settori connessi ma distinti del *knowledge management (KM)*, della *knowledge representation (KR)*, della *knowledge engineering* ecc. [1].

2 L'ordinamento sistematico

Il tipo più classico di *knowledge organization system (KOS)*, fin dal *Libro dei mutamenti* cinese e dall'albero di Porfirio neoplatonico, è la gerarchia di concetti progressivamente più specifici in ciascuna ramificazione.

Questa struttura è adottata molto spesso anche nei menù per la navigazione dei siti web, di solito in forme relativamente semplici che cercano di limitare il numero delle voci e dei gradi di specificità.

All'interno di ciascun ramo (*array*) della gerarchia, i concetti possono essere disposti (*sorted*) secondo diversi principî:

- in ordine di popolarità (per prime le voci più usate);
- in ordine alfabetico;
- in ordine sistematico.

Mentre i primi due principî sono adottati spesso dagli architetti degli informazione, il terzo sembra essere stato dimenticato dalla cultura informatica, nonostante fosse considerato fondamentale dai classici della KO.

Il grande biblioteconomo Ranganathan [2], infatti, definisce la classificazione anzitutto come una *successione conveniente* di concetti, dove “conveniente” significa aderente al più appropriato fra i seguenti principî generali:

- (1) cronologia,
- (2) evoluzione,
- (3) contiguità spaziale,
- (4) misura quantitativa,
- (5) complessità,
- (6) successione canonica,
- (7) garanzia bibliografica,
- (8) successione alfabetica.

Ciò che in biblioteconomia è chiamato “garanzia bibliografica” equivale grosso-modo alla popolarità, ossia a elencare per primi i concetti che ricorrono più spesso. Visto che gli otto principî sono elencati da Ranganathan esplicitamente in ordine di priorità, possiamo dunque notare che le due principali alternative all'ordine sistematico – la popolarità e l'alfabeto – sono considerate quelle meno utili, da applicare solo nell'impossibilità di adottare tutte le precedenti. Ad esempio, proporre i giorni della settimana in ordine alfabetico:

domenica
giovedì
lunedì
martedì
mercoledì
sabato
venerdì

è ovviamente meno efficace che elencarli nel loro ordine tradizionale (principio 6) o, se preferiamo, temporale (principio 1). Considerazioni simili valgono per altri tipi di concetti...

divorzio
fidanzamento
matrimonio
separazione

cascine isolate
città
metropoli
paesi

contrabbasso
viola
violino
violoncello

3 Dove hanno messo le donne?

Una decina di anni fa la redazione del MAI, un servizio di meta-ricerca simultanea nei cataloghi in rete delle biblioteche italiane, stava predisponendo l'interfaccia per la selezione dei cataloghi di singole regioni italiane [3].

I tecnici del CILEA misero a punto una cartina attivabile dell'Italia grazie alla quale, cliccando nell'area corrispondente a ciascuna regione, era possibile selezionarla per poi procedere con la ricerca. Tuttavia un altro redattore, Riccardo Ridi, osservò che il criterio della contiguità spaziale (il principio 3 di Ranganathan) sebbene ovvio per la maggior parte degli utenti italiani potrebbe risultare meno pratico per chi invece, come uno straniero o un utente poco ferrato in geografia, non si ricordasse bene che il Molise si trova fra l'Abruzzo e la Campania. Per questo motivo fu deciso di affiancare alla selezione geografica una possibilità alternativa di selezione tramite un elenco alfabetico delle regioni.

È chiaro che qui i due criteri, spaziale e alfabetico, svolgono funzioni complementari: un utente diverso, infatti, potrebbe ricordarsi visivamente in che parte della penisola si trovi Campobasso e però non rammentare il nome della suddivisione ammini-

strativa a cui la città è oggi assegnata (nell'Ottocento, tanto per dire, apparteneva al Regno delle Due Sicilie, che si troverebbe in tutt'altro punto dell'ordine alfabetico).

In questo senso, è noto che l'ordinamento sistematico offre all'utente non soltanto un indice, ma anche una guida che lo aiuta ad avvicinarsi gradualmente alla posizione di un concetto, anche quando non conosca il termine preciso con il quale esso viene indicato nel linguaggio del sistema. Se, per fare un piacere alla nostra vicina che tiene un pollaio, stiamo cercando chiarimenti sul comportamento alimentare delle donnole, di fronte ad un sito in inglese organizzato sistematicamente potremo cominciare ad esplorare le sezioni dedicate ai mammiferi carnivori, magari corredate di illustrazioni, fino a renderci conto che ciò che cerchiamo si trova alla voce *weasels*. Se fossimo stati obbligati a partire da una casella di ricerca per parole, senza conoscere la terminologia zoologica inglese, non ci saremmo probabilmente mai arrivati.

L'ordinamento sistematico dunque si abbina particolarmente bene alla funzione di *browsing*, ovvero di esplorazione di elenchi per scorrimento, che è notoriamente complementare al *searching* per parole.

Ma se è necessario e vantaggioso, perché l'ordinamento sistematico è tanto raro nei siti web, che quasi sempre adottano invece come KOS una *tassonomia* di voci si disposte in un albero gerarchico, ma elencate alfabeticamente? Il ritornello del venire incontro all'utente, che potrebbe non ricordarsi dove si trova il Molise, come abbiamo visto vale anche per invocare la soluzione opposta, a vantaggio di quelli che non sanno in quale regione si trovi Campobasso...

4 Il favoloso mondo della notazione

Un motivo importante per spiegare il sottoutilizzo dell'ordinamento sistematico è di natura tecnica. Infatti l'ordine alfabetico è spesso considerato una soluzione pratica perché i calcolatori lo conoscono già – è il loro criterio di ordinamento predefinito – e non è necessario insegnarglielo apposta. Semplicemente, si registrano le voci e si dà istruzione all'interfaccia di presentarle in ordine alfabetico.

Per ottenere un ordinamento sistematico, invece, occorre dare ai calcolatori (o ai bibliotecari che rimettono i libri al loro posto sugli scaffali) delle istruzioni aggiuntive. A questo scopo, ciascuna voce deve essere contrassegnata da un codice di posizione:

- B violini
- C viole
- D violoncelli
- E contrabbassi

- 1 lunedì
- 2 martedì
- 3 mercoledì
- 4 giovedì
- 5 venerdì
- 6 sabato
- 7 domenica

A questo punto, ciò che il calcolatore ordina non sarà più l'etichetta verbale di un concetto, ma il suo codice simbolico, espresso secondo una *notazione* convenzionale. Per questo stavolta scriviamo in carattere monospaziato i codici, che sono ora le unità di riferimento costituenti il nostro KOS, invece dei loro equivalenti verbali.

Le notazioni permettono di rappresentare e gestire non solo le voci di un array, ma anche le gerarchie:

- B violini
- BH violini barocchi
- C viole
- D violoncelli
- E contrabbassi

nonché altre strutture sofisticate dei sistemi di classificazione, come gli *ausiliari comuni* e le *faccette*, utilizzando vari ingegnosi dispositivi sintattici [4-5]:

- B violini
- B2D violini tedeschi
- B2I violini italiani
- BH violini barocchi
- BH2I violini barocchi italiani
- C viole
- C2I viole italiane

Se agli utenti la notazione offre una maggiore ricchezza informativa (non soltanto l'elenco bruto delle opzioni, ma anche le loro relazioni semantiche: relazioni d'ordine, gerarchiche, associative), agli sviluppatori di interfacce richiede però una maggiore complessità gestionale. Qualsiasi tabella di voci, infatti, dovrà prevedere un campo in più – la notazione associata all'etichetta verbale per quel concetto. Anche la generazione dinamica dell'interfaccia dovrà comprendere istruzioni (ad esempio un piccolo script PHP) per elaborare il contenuto del campo Notazione, utilizzandolo come criterio di ordinamento delle etichette corrispondenti.

È da notare che non necessariamente la notazione dev'essere visualizzata: per gli utenti, l'informazione più importante resta ovviamente l'etichetta verbale; la notazione è solo un meccanismo che permette di produrre ordinamenti più "intelligenti", e come tale può essere tenuta completamente dietro le quinte, o anche essere mostrata

per far intuire la logica di funzionamento del sistema, ma sempre senza distogliere eccessivamente l'attenzione dalla sostanza dei contenuti a cui fa riferimento.

5 Non solo menù: l'ordinamento nelle liste di risultati

Naturalmente, si potrebbe osservare che per una tassonomia che abbia soltanto poche voci in ciascun ramo, come succede nei menù di molti siti, non è poi così importante se queste voci siano ordinate sistematicamente, alfabeticamente o per popolarità: in ogni caso l'utente farà presto a scorrerle tutte e adocchiare quella che fa al caso suo.

Tuttavia l'ordinamento sistematico rivela appieno il suo potenziale quando si ha a che fare con quantità di voci molto maggiori. Nella tassonomia dei mammiferi (utilizzata per esempio in *Wikipedia*) ci sono molte ramificazioni, ciascuna con decine di gruppi animali imparentati fra loro: in questo caso i criteri per l'ordine di presentazione diventano più importanti.

Inoltre, anche quando le classi fra cui scegliere sono poche, ad essere molto numerose sono spesso le voci elencate in risposta alla ricerca, in fase di visualizzazione dei risultati. Ciascuna infatti sarà stata indicizzata con una o più classi dello schema adottato, e saranno presenti molte voci che condividono alcune o tutte le classi. È chiaro che in questo caso l'ordinamento adottato diventa importante, perché è noto che l'utente medio tende ad esaminare le voci una per una solo entro un certo limite quantitativo, il *futility point* [6].

Supponiamo per esempio che un insieme di link o di fotografie o di video dedicati agli strumenti ad arco siano stati indicizzati usando le classi dell'esempio precedente. La notazione, progettata in base a principi consolidati di organizzazione della conoscenza (come quello di Bliss che "il generale preceda lo specifico"), produrrà un ordinamento ben preciso – quello riportato sopra.

Come si può notare, i documenti sui violini in genere sono elencati prima di quelli sui violini tedeschi o italiani, in quanto queste ultime sono specificazioni che restringono l'ambito della trattazione: agli utenti viene offerto per primo il materiale generale e introduttivo, e successivamente quello più particolare. I documenti sui violini barocchi a loro volta vengono dopo, prima da soli e poi con le loro specificazioni geografiche, in quanto sono tipi più specifici della classe generale dei violini. Infine, esauriti tutti i tipi di violini disponibili, cominciano le viole – che li seguono in quanto vengono dopo nella classificazione musicale che va dai più acuti ai più gravi (principio 4).

Tutto questo, come dicevamo, può sembrare pedante finché ci limitiamo al nostro esempio con poche voci, ma diventa una preziosa fonte di ordine in liste di centinaia di voci, che sono molto frequenti nelle basi-dati di grandi negozi online, o nelle bibliografie specializzate, o nei cataloghi di materiali esposti in un grande museo. Per vedere un esempio italiano su grande scala si può provare il "navigatore Dewey" del catalogo della Biblioteca nazionale centrale di Firenze [7].

Al limite, prima ancora che la bontà dei criteri adottati, il fattore critico è che un qualche criterio ci sia, e che venga mantenuto con coerenza in tutto il sistema, in

modo che l'utente possa riconoscere un certo ordine e imparare a navigare secondo quello. Con le classificazioni a faccette è anche possibile (se le interfacce sono progettate tenendo conto della loro logica) variare secondo le preferenze dell'utente i criteri di ordinamento, così come quelli di ricerca [8].

6 Conclusione

Abbiamo visto come esistano diverse possibilità di ordinamento, la cui scelta può essere valutata a seconda delle esigenze pratiche. Le gerarchie offrono un primo strumento di organizzazione delle informazioni in tassonomie facilmente comprensibili e navigabili. Gli ordinamenti sistematici offrono ulteriori vantaggi, specialmente quando si abbia che fare con grandi quantità di voci estratte da una base-dati. D'altra parte essi richiedono la gestione, sia concettuale che tecnica, di una componente aggiuntiva: la notazione.

Gli strumenti per realizzarli sarebbero già disponibili: manuali di classificazione, sistemi di gestione di basi-dati online come MySQL, linguaggi di script capaci di manipolare stringhe e produrne contenuti web come PHP. Quello che ancora manca è probabilmente una maggiore diffusione di applicazioni, simili a quelle già esistenti, che fungano da esempio aiutando a far crescere la consapevolezza delle funzioni e delle potenzialità dell'ordinamento.

Riferimenti bibliografici

1. Gnoli, C., Marino, V., Rosati, L.: *Organizzare la conoscenza: dalle biblioteche all'architettura dell'informazione per il Web*. Hops-Tecniche nuove, Milano (2006)
2. Ranganathan, S.R.: *Prolegomena to Library Classification*. 3rd ed. SRELS, Bangalore (1967), part F. Ed. dig. in UA Campus Repository, University of Arizona, arizona.openrepository.com/arizona/bitstream/10150/106370/7/ProlegomenaF.pdf
3. *Metaopac Azalai italiano, Ricerca per regione*, www.aib.it/aib/opac/mai3.htm3 (2001-)
4. Vickery, B.C.: *Notational symbols in classification*. *J. Doc.* 8, 14-32 (1952); 12, 73-87 (1956); 13, 72-77 (1957); 14, 1-11 (1958)
5. Gnoli, C.: *L'alfabeto e la sindrome di Sariette*. *AIDA informazioni*, 24, 3-4 (2006), aidainformazioni.it/aidainformazioni.it/pub/gnoli342006.html
6. Blair, D.C.: *Searching biases in large interactive document retrieval systems*. *J. Am. Soc. Info. Sci.* 31, 271-277 (1980)
7. *Biblioteca nazionale centrale di Firenze*, www.bncf.firenze.sbn.it (1999-), *Catalogo in linea (OPAC) del polo BNCF*
8. Rosati, L.: *Le faccette in architettura dell'informazione*. In: Barazia, C., Gnoli, C. (cur.). *Le dimensioni dell'informazione*. Documenti ISKO Italia, www.iskoi.org/doc/dimensioni3.htm (2006-)