
La *classificazione a faccette* fra Knowledge Management e Information Architecture (parte I)

Luca Rosati
(<http://www.lucarosati.it/>)

Il Concetto di Classificazione a Faccette

Per poter trasformare l'informazione in conoscenza è necessario anzitutto assicurare un accesso **flessibile** e **multi-dimensionale** alla informazione stessa. I contesti aziendali attuali sono caratterizzati infatti, oltretutto da una grande mole informativa, anche da rapidi mutamenti del modo di produrre e organizzare tale informazione.

Per far fronte a questo scenario (non solo all'interno di sistemi di knowledge management propriamente detti ma anche, più semplicemente, in siti, intranet e data-repository di qualsiasi genere) è necessario progettare e realizzare sistemi di classificazione e accesso all'informazione **aperti, flessibili** e **adattivi**, in grado cioè di adattarsi a una pluralità di esigenze di archiviazione e/o interrogazione.

A questo scopo, negli ultimi anni, si sono intessute fitte relazioni fra discipline come l'architettura dell'informazione (che si interessa dell'organizzazione dell'informazione all'interno di sistemi digitali), il Knowledge Management e la biblioteconomia (la stessa architettura dell'informazione nasce da una applicazione al web della biblioteconomia classica).

La tesi sicuramente più accattivante [...] è quella di una interdisciplinarietà progressiva e marcata tra i diversi settori della conoscenza, a conferma di una eterna circolarità del flusso informativo e in adesione ai moderni concetti di rete, integrazione e condivisione dei dati. In questa dimensione intercomunicante di esperienze conoscitive, non si rivela dunque azzardato ipotizzare accostamenti tra discipline quali biotecnologia, astrofisica e geopolitica, ad esempio, aree di studio che solo una tradizione obsoleta perdura a confinare entro ambiti definiti privi di connessioni
(E. Poltronieri, [Bollettino AIB 1997 n. 1 p. 104-106](#)).

In un simile scenario, hanno così conosciuto una nuova primavera le nozioni di "**faccette**" e di "**classificazione a faccette**" (o "classificazione multidimensionale"), elaborati nell'ambito della biblioteconomia classica – da Ranganathan in avanti, con un importante incremento negli anni '60-'70). La classificazione a faccette (in inglese *Faceted Classification*) è un metodo di classificazione dell'informazione il cui valore distintivo consiste appunto nell'essere un sistema aperto e adattivo.

La classificazione mediante faccette sfrutta un sistema di attributi (**metadati**) mutuamente esclusivi rappresentanti ciascuno un aspetto o proprietà persistente dell'oggetto e capaci – nel loro insieme – di descrivere esaustivamente l'oggetto stesso. Tali attributi sono detti **faccette** (termine introdotto in questa accezione dal biblioteconomista e matematico indiano Ranganathan), e sono contraddistinti da queste peculiarità:

- sono invariabili dal punto di vista semantico (ad es. la proprietà COLORE di un oggetto può variare in termini di valori che può assumere – giallo, rosso etc. – ma è invariabile come concetto; cioè quell'oggetto avrà sempre un colore)
- costituiscono un insieme aperto, per cui è sempre possibile aggiungere nuove faccette a quelle già esistenti
- sono utilizzabili come attributi di ricerca sia singolarmente sia in combinazione.

Tali caratteristiche rendono particolarmente efficace l'adozione di questo sistema in ambienti digitali, per un più veloce ed efficiente ritrovamento dell'informazione (cf. Wynar, cit. da P. Murray, [Faceted classification of information](#)).

Funzionamento

Per fare un esempio concreto, poniamo di dover costruire un catalogo di vini. Potremmo allora rappresentare ciascun vino mediante questo set di proprietà o **faccette**:

- COLORE (C)
- VITIGNO (V)
- PROVENIENZA (P)
- ANNATA (A)
- ...

Tali faccette sono mutuamente esclusive l'una rispetto alle altre, vale a dire che fra esse non vi è sovrapposizione semantica (ciò non significa però che un vino può comparire solo in una di esse; anzi, ogni vino è rappresentato dall'insieme delle varie faccette).

Ciascuna faccetta, potrà essere a sua volta suddivisa/scomposta in più **topic** (o foci, o faccette di secondo livello – la nomenclatura può variare da autore ad autore):

- COLORE (C)
 - rosso (r)
 - bianco (b)
 - rosato (s)
- VITIGNO (V)
 - aglianico (agl)
 - merlot (mr)
 - sangiovese (sg)
 - ...

- PROVENIENZA (P)
 - campania (c)
 - piemonte (p)
 - sicilia (s)
 - ...
- ANNATA (A)
 - 1961
 - ...

Una bottiglia di *Taurasi 1996*, risulta descritta da queste coordinate: C(r), V(agl), P(c), A(96).

Ciò significa che:

- è possibile ricercare questo vino utilizzando una sola delle sue faccette, oppure una combinazione di esse (es. C(r) & V(agl) & P(c))
- è possibile in ogni momento implementare nuove faccette di primo livello (ad es. PRODUTTORE, ABBINAMENTI, GRADAZIONE etc.)
- è possibile in ogni momento aumentare il numero dei topics o faccette di secondo livello interni a una faccetta di primo livello.

Classificazione a Faccette vs. Classificazione Tradizionale

Nei sistemi di classificazione tradizionali (detti anche *tassonomie* tradizionali o sistemi **enumerativi**), ogni elemento è classificato sotto una e una sola *categoria*. Esso possiede una corretta e **univoca** collocazione all'interno di un unico schema, ampio e gerarchicamente molto profondo, e può essere ritrovato attraverso un percorso a gradini categoria padre > categoria figlio. ([Giovanni M. Sacco](#) definisce le tassonomie tradizionali come dei *codificatori di proprietà* che, partendo dalla categoria-padre generano via via categorie-figlie mediante l'aggiunta di nuove proprietà).

Ne sono un esempio il *sistema di Linneo*, il *sistema decimale Dewey* e quello della *Library of Congress* (va tuttavia osservato che i sistemi Dewey e Library of Congress, dopo il 1960 – sulla scorta dell'influenza di Ranganathan, furono ritoccati accogliendo, se pure limitatamente, parte della logica delle faccette).

Un sistema di questo tipo è quindi **monodimensionale** (il criterio di catalogazione è unico), e molto esteso in **verticale** – seppure è possibile cercare al suo interno attraverso più modalità (comunque limitate: essenzialmente, titolo, soggetto, autore).

Essi costituiscono una sorta di 'recipienti' chiusi, disposti in sequenza, e concepiti in forma di matrioska (sistema di scatole nella scatole, con forte gerarchizzazione).

Alla verticalità dei sistemi di catalogazione tradizionali (i. e. eccessiva ramificazione in profondità delle gerarchie), e alla loro rigidità, la classificazione a faccette contrappone un sistema di classi (*faccette*) **orizzontale** e **aperto** (laddove ciascuna faccetta è descrittiva di una proprietà o faccia dell'oggetto).

I principali tratti distintivi di un sistema a faccette sono così sintetizzabili:

- **pluridimensionalità**: inversamente ai sistemi tradizionali, nei sistemi a faccette, ogni oggetto è classificato secondo una pluralità di attributi (detti – appunto – faccette)
- **persistenza**: tali attributi/faccette costituiscono proprietà essenziali e persistenti dell'oggetto; in questo modo l'impatto (sullo schema di classificazione) di eventuali cambiamenti (di nomenclatura, di workflow etc.) è fortemente ridotto o nullo
- **scalarità**: è sempre possibile aggiungere una nuova faccetta descrittiva di un nuovo aspetto dell'oggetto
- **flessibilità**: esiste una pluralità di chiavi di accesso parallele (*faccette*); ogni oggetto può essere reperito utilizzando un singolo attributo di ricerca (o *faccetta*) alla volta, oppure più attributi insieme in combinazione
- non è necessario conoscere il nome della/e categoria/e in cui l'oggetto è collocato (in un contesto di rapido mutamento questo costituisce un indubbio vantaggio)
- il criterio di ordinamento non è necessariamente gerarchico (dal generale al particolare, dall'intero alla parte etc.); esso può anche essere alfabetico
- è possibile costruire relazioni persistenti fra le diverse faccette, tali da fornire una rappresentazione della conoscenza contenuta nel sistema stesso: ad es. Mario Rossi (faccetta PERSONA) < è un impiegato della > Azienda Z (faccetta AZIENDA) e < sviluppa > Applicazioni XML (faccetta PRODOTTI).

Dati alla mano, le applicazioni basate su sistemi di classificazione a faccette hanno dimostrato di essere molto più veloci da consultare e di garantire più alte percentuali di successo rispetto a quelle più tradizionali. Giovanni M. Sacco (Università di Torino) ha posto sperimentalmente a confronto due sistemi di classificazione, uno basato su **tassonomie tradizionali**, l'altro su **tassonomie dinamiche** (*faccette*): eseguendo una ricerca su un corpus di 114.042 documenti descritti da 1100 concetti (dove ogni documento è classificato in media attraverso 11.19 topics), la media dei risultati è stata

- di 1246.22 documenti con tassonomie tradizionali (in tal caso l'utente deve ispezionare manualmente una tale mole di documenti – nessun utente lo fa)
- di 27.68 documenti con tassonomie dinamiche, dopo aver operato un singolo 'zoom'.

In conclusion, whereas a conventional taxonomy would require the user to inspect result lists of several hundreds documents, a dynamic taxonomy requires only one zoom operation to produce an average result size of 13 to 27 documents
(G. M. Sacco, [Conventional Taxonomies vs. Dynamic Taxonomies](#), p. 10).

Proprio per queste sue peculiarità la classificazione a faccette è stata definita da Eugenio Gatto (uno studioso di Ranganathan) anche come **classificazione multidimensionale** (cf. [intervento di Claudio Gnoli](#) nella mailing list “Faceted Classification Discussion”).

Un eccellente esempio della scalabilità e flessibilità di questo sistema di classificazione è offerto da [Boxes and Arrows](#), rivista online dedicata alla architettura dell’informazione.

Dalla data della sua prima apparizione (marzo 2002) fino a settembre 2002 il sito presentava un solo criterio di archiviazione degli articoli: quello per data; la barra di navigazione principale presentava così due sole opzioni, le *tab* “Current” e “Previously”.

A partire da ottobre 2002, la barra si è arricchita di 2 nuove opzioni, le *tab* “Categories” e “Authors”, corrispondenti a due nuovi criteri di classificazione degli articoli:

- per argomento (*Categories*)
- per autore (*Authors*).

Ciò è reso possibile dal fatto che la rivista (e il content management system che la alimenta) poggiano su un sistema di indicizzazione a faccette, che prevede l’associazione ad ogni articolo di:

- una data
- un autore
- una o più categorie/argomenti
- un set di parole chiave.

Questo permette di accedere agli articoli mediante una pluralità di logiche, soddisfacendo così una uguale pluralità di esigenze e scopi di navigazione.

Non è escluso ad esempio che, in un prossimo futuro, possa comparire anche un *indice analitico* dei contenuti (costruito sulla scorta delle parole chiave associate ai vari articoli) - a ulteriore dimostrazione della flessibilità di questa struttura.

Anche i portali come Yahoo offrono un esempio di questo tipo di approccio, sebbene applicato in una forma spuria (cioè non così come indicato dalle scienze librarie).

Classificazione a Faccette e Gestione della Conoscenza

Guardare alla classificazione a faccette esclusivamente come a un apparato teorico coniato dalle scienze librarie è quindi limitativo. Questo approccio, infatti, costituisce la formalizzazione di una tecnica di comunicazione che noi spesso utilizziamo in un ampio spettro di contesti, dalla organizzazione delle informazioni personali (specie in agende elettroniche o nei bookmarks) fino ai giochi (Dire, fare, baciare; *Twenty Question*).

La stessa metodologia seguita in molti cataloghi o indici analitici cartacei riflette questa logica di fondo (in maniera più o meno ibrida), anche se in pochi forse definirebbero l’indicizzazione come uno schema di classificazione.

Questo perché la classificazione a faccette possiede importanti vantaggi rispetto ad altri sistemi.

In particolare, sono soprattutto le caratteristiche indicate al paragrafo precedente con le etichette di **pluridimensionalità**, **persistenza**, **scalarità** e **flessibilità** a rendere questo sistema di catalogazione particolarmente vantaggioso per le organizzazioni aziendali soggette a forte dinamismo e cambiamento.

Tali caratteristiche, infatti, prevengono il degrado dei repository aziendali in quanto evitano che i cambiamenti si ripercuotano negativamente sulla organizzazione dell'informazione.

Phil Murray (di KMConnection) nel suo [Faceted Classification of Information](#) sintetizza così i vantaggi della adozione di un sistema di classificazione a faccette per la gestione della conoscenza aziendale:

- pur non conoscendo il nome o la collocazione di un oggetto, è possibile nondimeno renderlo comprensibile e raggiungerlo, descrivendolo attraverso un insieme di categorie (faccette) mutuamente esclusive (nell'esempio del vino al par. 2: COLORE; VITIGNO; PROVENIENZA etc.)
- possibilità di aggiungere nuove faccette a quelle già stabilite in una fase precedente (carattere aperto del sistema)
- possibilità di combinare fra loro i valori delle diverse faccette
 - per fornire una descrizione più precisa dell'oggetto
 - per limitare i risultati di una ricerca
 - per descrivere relazioni esistenti fra vari oggetti (rappresentazione della conoscenza contenuta nel sistema).

Applicazioni

Nel prossimo articolo presenterò una rassegna di alcune applicazioni della classificazione a faccette in:

- siti web
- sistemi generali di catalogazione online
- linguaggi di marcatori (XML derivati) per la creazione e lo scambio di sistemi di classificazione a faccette.

Per Approfondire

In italiano:

- Claudio Gnoli, [La classificazione a faccette](#)
- Elisabetta Poltronieri, [Bollettino AIB 1997 n. 1 p. 104-106](#)
- Luca Rosati – Information Architecture, [risorse su Faccette e metadati](#).

In inglese:

- Phil Murray, [Faceted classification of information](#)
- [Faceted Classification Discussion](#)
- Giovanni M. Sacco, [Conventional Taxonomies vs. Dynamic Taxonomies](#).