



**INDIRIZZO GRAFICO
AREA DI PROFESSIONALIZZAZIONE**

a.s.: 2006 / 2007

**ESPERTO DI DESKTOP PUBLISHING
E WEB/FLASH GRAPHIC DESIGN**

L'architettura del sito Web

prof. Aldo Gorla

Guida Architettura dell'Informazione

Introduzione

1. **Perché l'architettura dell'informazione,**
L'architettura dell'informazione tra usabilità e accessibilità
2. **I limiti dell'usabilità come metodo di progettazione**
I problemi dell'eccessivo accumulo di informazioni e della loro archiviazione
3. **A cosa serve l'architettura delle informazioni**
Scripta manent, verba volant, digitalia mutant: controllare le mutazioni
4. **Che cosa è l'architettura delle informazioni**
Breve storia e applicabilità quotidiana dell'Architettura dell'informazione
5. **I vantaggi dell'Architettura dell'informazione**
Le opportunità che offre una buona architettura dell'informazione
Bad Architecture, ovvero quattro esempi da non seguire
6. **Motori di ricerca e rilevanza delle informazioni**
Conciliare i risultati dei motori di ricerca con la rilevanza delle informazioni
7. **Le gabbie iniziali non bastano più**
Progettare un sito con l'occhio rivolto al futuro accumulo all'informazione
8. **Imputato Nielsen, si alzi!**
L'accumulo di informazioni nel sito di Jacob Nielsen
9. **I blog fanno venire il mal di testa**
L'organizzazione dell'informazione in un qualsiasi weblog
L'organizzazione dei contenuti
10. **L'"Information Brokering"**
Affrontare il problema della reperibilità dell'informazione
I principi dell'information brokering
Come identificare e rendere evidenti i contenuti del proprio sito Web
11. **La schematizzazione dei contenuti**
12. **Gli schemi percettivi e gli schemi esatti**
Le modalità di schematizzazione dei contenuti all'interno dei siti web - I
13. **Gli schemi ambigui**
Le modalità di schematizzazione dei contenuti all'interno dei siti web - II
14. **Come classificare: card sorting cartaceo**
Utilizzare il metodo del card sorting attraverso cartoncini
15. **Come classificare: card sorting digitale**
IBM EZSort: un programma di card sorting digitali
16. **Dopo il Card Sorting: il dilemma delle categorie**
Come strutturare e classificare le categorie dopo averle organizzate
17. **Architettura dei task e timone dell'homepage**
Prevedere le azioni dell'utente e gestire le aree dell'homepage
Architettura dei contenuti: progettare un CMS
18. **Il Content Management System e i metadati**
L'importanza dei metadati all'interno di un sistema di gestione dei contenuti
Architettura dei contenuti: progettare un CMS
19. **Metadati di categoria**
Organizzare le categorie dei nostri contenuti
20. **Metadati di tempo**
Organizzare il tempo di pubblicazione dei nostri contenuti
21. **Metadati linguistici e redazionali**
Definire la lingua e i diversi micro-contenuti delle nostre pagine

22. **. Metadati di authoring e di target**
Definire chi scrive e chi leggerà i nostri contenuti
23. **Metadati per il Thesaurio**
Correlare contenuti all'interno del nostro sito
24. **Thesauri complessi**
Gestire associazioni complesse per favorire la navigazione
Architettura delle Informazioni e navigazione
25. **I sistemi di navigazione**
Le strutture di navigazione all'interno di un sito Web
L'Architettura delle Informazioni nel motore (di ricerca)
26. **Analisi dei termini della ricerca**
Come far analizzare dal motore i termini inseriti
27. **La ricerca**
Come varia la ricerca in base alle parole utilizzate
28. **Metadati e presentazione**
Come presentare i risultati e rendere migliori le ricerche
29. **I visitatori-autori: tecniche di filtraggio collaborativo**
Come l'analisi della navigazione degli utenti ci aiuta a raggruppare i contenuti
30. **. Esempi di filtraggio collaborativo**
Amazon e stock xchg: come farsi aiutare dai visitatori

Perché l'architettura dell'informazione

Quando si parla di **usabilità** si pensa quasi sempre alla percezione del layout, o, per dirla più semplicemente, all'aspetto grafico: colore dei link, posizione della barre, animazioni, peso delle immagini.

Mentre quando si parla di **accessibilità**, si toccano per lo più aspetti tecnici: tempi di caricamento, possibilità di visualizzazione per gli utenti diversamente abili, diversificazione della presentazioni dei contenuti per tutti i tipi di supporto tecnologico.

In questo senso, l'accessibilità è diventata un requisito indispensabile per tutti i siti delle pubbliche amministrazioni, e a ragione: perché tutti i visitatori devono poter accedere alle informazioni di pubblica utilità, e perché un sito accessibile è comunque un sito meglio costruito anche per un utente abile.

Esiste **un altro tipo di inaccessibilità**, però, e si annida anche in siti costruiti secondo le più aggiornate linee guida tecniche. In molti portali, soprattutto di enti pubblici, si riscontra un aumento esponenziale delle informazioni pubblicate: più o meno indispensabili, più o meno recenti, dedicate a diversi tipi di pubblico, che convivono tutte insieme nello stesso ambiente, determinando così un sovraccarico, una sorta di "inquinamento da informazioni" che rende spesso impossibile arrivare al contenuto davvero rilevante per l'utente.

All'**accumulo indiscriminato di informazioni** va contrapposta una gestione ragionata dei contenuti, che non significa scrivere e pubblicare meno, ma ordinare, catalogare, filtrare, richiamare all'occorrenza tutti i contenuti collegati, in altre parole: un'ecologia delle informazioni.

L'Information Architecture (Architettura dell'informazione) è così la disciplina che permette di contrastare gli effetti negativi dell'accumulo di informazioni e di andare verso un'ecologia dei sistemi complessi di contenuti.

I limiti dell'usabilità come metodo di progettazione

Un sito progettato da zero contiene un numero limitato di informazioni.

Anche il migliore dei layout, in grado di superare tutti i test di usabilità, alla fine **capitola davanti al costante accumulo di notizie**: progettato per una visione statica o fotografica di un sito infatti, non tiene conto del divenire. O meglio: anche un layout espandibile cede di fronte all'accumularsi delle informazioni e difficilmente gli utenti arriveranno a cliccare fino alla quindicesima pagina di archivio per trovare qualcosa che potrebbe interessarli.

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

In altre parole, nel momento in cui un contenuto affonda nel mare magnum del passato va praticamente perduto (a meno che non venga recuperato dagli spider di un motore di ricerca, ma di questo di parlerà in seguito).

Che cosa accade quando la quantità di contenuti di un sito web diventano considerevoli?

- Il numero delle pagine di archivio aumenta, e già dopo il terzo click spesso l'utente si demotiva;
- Il numero delle opzioni tra cui scegliere aumenta, ma in pratica si riduce la reale capacità di scelta (cercare, analizzare e scegliere tra 20 non è come cercare, analizzare e scegliere tra 100).
- I motori di ricerca interni al sito non riescono a filtrare e a selezionare un numero gestibile di contenuti che siano anche effettivamente rilevanti rispetto alla parola chiave cercata.
- Davanti alla mascherina bianca di un motore di ricerca non sempre i visitatori hanno ben chiaro che cosa stanno cercando.

Alcuni siti aggirano il problema dell'**overload** rinunciando a tenere online gli archivi, con il rischio di perdere riferimenti al passato che possono sempre rivelarsi preziosi.

Dal lato opposto è molto interessante la scelta di Repubblica.it: spesso accanto a una notizia propone una serie di link dall'archivio che ricostruiscono gli antefatti e il contesto esterno del pezzo che si sta consultando.

Figura 1

Il ministro degli Esteri è un ragazzo svedese di
con precedenti penali e problemi psichici

er la polizia Lindh è risolto

o del giovane arrestato due giorni fa

Il caso è
lizia svedese
ministro degli
Lindh, accoltellata
corso in un
o di
giovane di 24
ovic,
e jugoslava,
merosi
e un diversi
iatrici.

anni fa, in
a liberazione **Anna Lindh**

ragionato,
ascoltato oggi dal giudice, che ha ritenuto le
ora sufficienti a prolungare la detenzione
tre due settimane. Dopo aver rilasciato quello
era sembrato il maggiore sospetto, stavolta la
apere che gli indizi raccolti sono "più gravi" e
resi accusatoria non solo "ragionevole" ma

esta quindi in carcere, mentre è stata
lizia per verificare le condizioni psichiche del
ebbe un carattere instabile e violento. Gli
dunque altri quindici giorni, cioè fino al 10

**è in ec
con Rep**

Invia questo articolo

La notizia

Lo storico della notizia

**Svezia, per la polizia
il caso Lindh è risolto**

**L'ARCHIVIO
pubblica.it**

**Stoccolma, arrestato
presunto killer della
Lindh**

**Svezia, è morta
Anna Lindh**

**Stoccolma, accoltellata
il ministro degli Esteri**

Stampa questo articolo

Il sito di Repubblica.it, organizza manualmente i collegamenti all'archivio, sfruttando nel web l'abitudine giornalistica di archiviare tutti gli articoli in una miriade di cartelle e sottocartelle (invisibili all'utente) in modo da avere sempre ricostruito lo storico di una notizia.

A cosa serve l'architettura delle informazioni

La scrittura su carta in alcuni casi rimane anche troppo **fedele a se stessa**. Quante volte chi lavora nella comunicazione si è sentito rispondere che è troppo tardi per una modifica al programma degli incontri, perché la brochure è già andata in stampa. Lo stesso capita a un libro che fresco di tipografia manca di un riferimento attuale nelle note e nella bibliografia.

Nel bene e nel male il digitale è a metà tra scritto e orale, il **digitale muta**. La forza della pagina digitale è quindi la sua possibilità di conservare (scripta manent) ma allo stesso tempo di mutare, per adattarsi alle varie forme del divenire. E il contesto varia. L'articolo scritto su carta sta fermo, (scripta manent) rimane se stesso. E non può essere diversamente da così. Ma l'universo potenziale dei suoi riferimenti prosegue inesorabilmente.

L'**architettura delle informazioni** permette di tenere immutato l'articolo (scripta manent) e contemporaneamente di aggiornare continuamente i suoi riferimenti all'esterno che inevitabilmente si muove (digitalia mutant).

Perché l'architettura delle informazioni e non il web semantico

Le parole chiave e le classificazioni servono a questo: a identificare alcuni concetti chiave facendo in modo che esse rendano il contenuto rintracciabile e lo aggancino a riferimenti correlati.

Il **web semantico** - ammesso che funzioni - questo non può sempre farlo. Infiniti percorsi narrativi, rivoli e microstorie, neologismi e instant words si dispiegano lungo l'accumularsi delle notizie, e il web semantico non può ricostruirli tutti a priori.

Non solo. Ponendosi come punto di riferimento internazionale, può rischiare di naufragare nell'annoso problema delle differenze linguistiche e culturali. Non tutti i termini sono traducibili da una lingua all'altra, non tutte le culture schematizzano allo stesso modo e riescono tranquillamente a importare l'una nell'altra gli schemi prescelti. (chi non ricorda i dodici modi di classificare la neve della popolazione Inuit? da: "Il senso di Smilla per la Neve")

Architettura delle informazioni in tutte le direzioni

Analizzando i dati di accesso di portali ad alto traffico ci siamo accorti che una percentuale sempre più grande di visitatori non entra dalla home page del sito, bensì **dai motori di ricerca**. E i motori di ricerca li portano direttamente alle foglie. Gli utenti giungono così immediatamente all'ultimo elemento della classificazione, quello che d'ora in poi verrà chiamato **unità di contenuto**, senza passare dal via. Quanto più saremo in grado di arricchire l'esperienza della foglia, tanto più nostri visitatori si sentiranno spinti a visitare anche i rami dell'albero e tanto più saranno attratti dalle associazioni a cui quell'unità di contenuto rimanda.

Che cosa è l'architettura delle informazioni

Gli esseri umani hanno sempre avuto la necessità di organizzare e strutturare le informazioni da loro stessi prodotte. Accanto alla proliferazione della parola, si è sviluppata una serie di tecniche per arginare il possibile caos. Già nel 330 avanti Cristo, la biblioteca di Alessandria era dotata di un catalogo composto da 120 rotoli di pergamena. (Cfr. Architettura delle Informazioni per il World Wide Web)

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

Nel **1873**, Melville Dewey ha sviluppato il sistema decimale che oggi porta il suo nome, e che consiste in uno strumento per organizzare e fornire accesso al crescente numero di libri prodotti dalla cultura del suo tempo e da quella precedente.

Nel **1930**, il matematico e biblioteconomista indiano Ranganathan elabora un nuovo sistema di classificazione: faceted Classification (in italiano "Classificazione a faccette"). Negli anni '**50-'60**, il Classification Research Group (CRG) riprende e affina la teoria di Ranganathan. Nell'**ultimo decennio**, Rosenfeld e Morville con il loro Information Architecture per the world wide web hanno ripreso e portato avanti il lavoro di Ranganathan per adattarlo al web.

L'architettura delle informazioni nella vita quotidiana

Molti di noi, attraverso l'esperienza di libri e **biblioteche**, hanno acquisito familiarità con i concetti di base dell'organizzazione delle informazioni.

In una bancarella di libri usati, accatastati alla rinfusa, è divertente **esplorare l'ignoto** polveroso andando a caccia di preziose rarità editoriali sopravvissute a oscuri cataclismi. Pensiamo invece ad una intera biblioteca di 10.000 libri non organizzata: non ospita i libri, li nasconde! Il miglior modo per nascondere un libro in una biblioteca è spostarlo in un altro scaffale. Questa è la tecnica usata da molti studenti per conservare da un giorno all'altro il privilegio di accesso a un testo d'esame ...

Il **catalogo di una biblioteca** non è l'equivalente di un elenco telefonico dei libri, non vi sono solamente libri ordinati per autore, data, editore, ecc. Esistono **criteri ragionati**, la classificazione per soggetti, o il codice Dewey appunto, che organizza i titoli per aree tematiche, in alcuni casi forse arbitrariamente, ma in maniera estremamente utile per chi si trova in fase di esplorazione tra gli scaffali.

Anzi: quante volte, cercando qualcosa, per una misteriosa catena di associazioni, siamo approdati ad altre informazioni che non stavamo cercando ma che si rivelano la risposta preziosa a un quesito in sospenso da tempo? In inglese questa esperienza si chiama **serendipity**. Cercare un argomento all'interno di un sistema ben organizzato è già apprendimento.

L'architettura delle informazioni "dà una mano" al caso. Applicata ai siti web, **rafforza i sistemi di navigazione**: organizza i concetti per categorie e rende possibile la creazione dinamica di percorsi associativi estemporanei.

I vantaggi dell'Architettura dell'informazione

Ma quali sono le opportunità concrete che questo approccio ci fornisce? Eccole:

- Una buona architettura delle informazioni **costruisce schemi e sistemi di navigazione fissi** (elenchi, categorie e sottocategorie) in cui fare rientrare ogni elemento. Ha un metodo per organizzare gli schemi, non si affida al caso. Sa che esistono schemi oggettivi (come gli indici alfabetici) che sono precisi, ma non aggiungono nulla a quello che già si sa, e schemi ambigui, le categorie tematiche (per esempio il codice Dewey) che proprio perché sono ambigui aggiungono conoscenza.
- Una buona architettura delle informazioni **identifica thesauri e associa metadati ad ogni contenuto del sito**. Quando due pagine sono associate a uno stesso termine di un vocabolario controllato, allora potranno essere collegate tra loro tramite un link, anche se non appartengono alla stessa categoria di navigazione. Il vocabolario controllato può essere composto anche da migliaia di termini, e gli effetti delle connessioni possono arrivare dopo molto tempo, cioè dopo che le informazioni si sono accumulate (e gli schemi di navigazione iniziali iniziano a essere affollati). E' un po' come se gli scaffali di una biblioteca si mettessero a volare, abbandonassero le impostazioni del codice Dewey e si riorganizzassero attorno a un termine del

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

thesauro. Gli scaffali delle biblioteche, si sa, sono piuttosto rigidi e non possono volare; ma i database non sono rigidi, se ben strutturati ci restituiscono le informazioni in scaffali sempre ordinabili con logiche differenti.

- Una buona architettura delle informazioni **sa anche prevedere metadati aperti**. Un sistema di informazioni, soprattutto in campo giornalistico, non è un mondo chiuso: vede nascere continuamente filoni tematici e neologismi. Non si può in questi casi prendere un vocabolario tematico vastissimo (ma bloccato) e incastonarlo nel sito una volta per tutte. Occorre pensare a un sistema che accetti continuamente parole, nuovi termini, e che sia in grado di correlare tra loro le notizie associate a questi termini. Che vada a recuperare le notizie associate negli archivi, anche se nella classificazione cronologica si trovano nella cinquantesima pagina, e riportarle lì, nel presente di quella pagina che il visitatore sta leggendo, perché di quella notizia costituiscono il radicamento con il passato e con il presente: in una parola, il contesto.
- Una buona architettura delle informazioni **tiene conto dei comportamenti dei visitatori**. Vede che se i visitatori passano continuamente da un punto all'altro del sistema, significa che tra questi due punti esiste un legame molto stretto, anche se tra gli autori nessuno se n'era accorto. Una buona architettura delle informazioni sa rendere i visitatori anche un po' autori.

Gli effetti di una buona architettura delle informazioni si ritrovano concretamente nell'arricchimento dell'esperienza della navigazione e nell'efficienza dei motori di ricerca interni al sito.

Motori di ricerca e rilevanza delle informazioni

Proviamo a ricercare la parola "lavoro" nel motore di ricerca all'interno del pluripremiato portale della Regione Emilia-Romagna. Ci accorgeremo che, tra le risposte, uno dei risultati più importanti (Agenzia Emilia Romagna Lavoro) è nascosto tra una serie di mozioni e interventi in cui la parola "lavoro" compare, ma non ha la portata generale che ci si aspetterebbe (la maggior parte dei link è infatti dedicata ai problemi degli insegnanti precari).

Questo accade perché l'articolo in questione contiene sì la parola lavoro, ma non nell'accezione generale, e il portale nel suo complesso manca di un sistema di mappatura per parole chiave. È vero che tra gli argomenti predefiniti il sito della regione E-R. contiene la categoria "Lavoro", ma un utente dovrebbe essere in grado di raggiungere gli stessi contenuti sia utilizzando i link, sia digitando la parola che lo interessa nel motore di ricerca interno.

Figura 2



Un altro significativo esempio: La pagina "Diritto allo studio" esiste nel portale dell'Università di Bologna, solo che facendo una ricerca testuale non emerge nessun risultato.

Figura 3



Le gabbie iniziali non bastano più

tra i contenuti di un sito in continuo aggiornamento, intervengono nuovi tipi di relazioni che non potevano essere previsti fin dall'inizio. Il sito per quanto ben progettato per crescere nel layout e per tutto ciò che riguarda il design, non è più progettato per gestire l'accumulo delle informazioni.

L'usabilità ci indica come essere credibili, come rafforzare l'interazione uomo macchina, ma non è in grado verificare e promuovere le diverse relazioni che nascono fra i contenuti.

Vediamo un esempio di notizie accumulate in maniera indiscriminata, in un sito che in partenza poteva essere usabile:
<http://www.umtsarea.com/archive.php>.

Abbiamo un lungo elenco di news che si giustappongono l'una all'altra.

Figura 4



Le informazioni, giustamente, sono circostanziate con data e ora; per renderle più fruibili, sarebbe consigliabile riunirle in sottogruppi e organizzarle in percorsi (anche temporanei) che si creano attorno a un argomento. Il design mira a organizzare contenuti stabili, ma nel mare magnum delle notizie, spesso si formano e dovrebbero affiorare percorsi tematici.

Imputato Nielsen, si alzi!

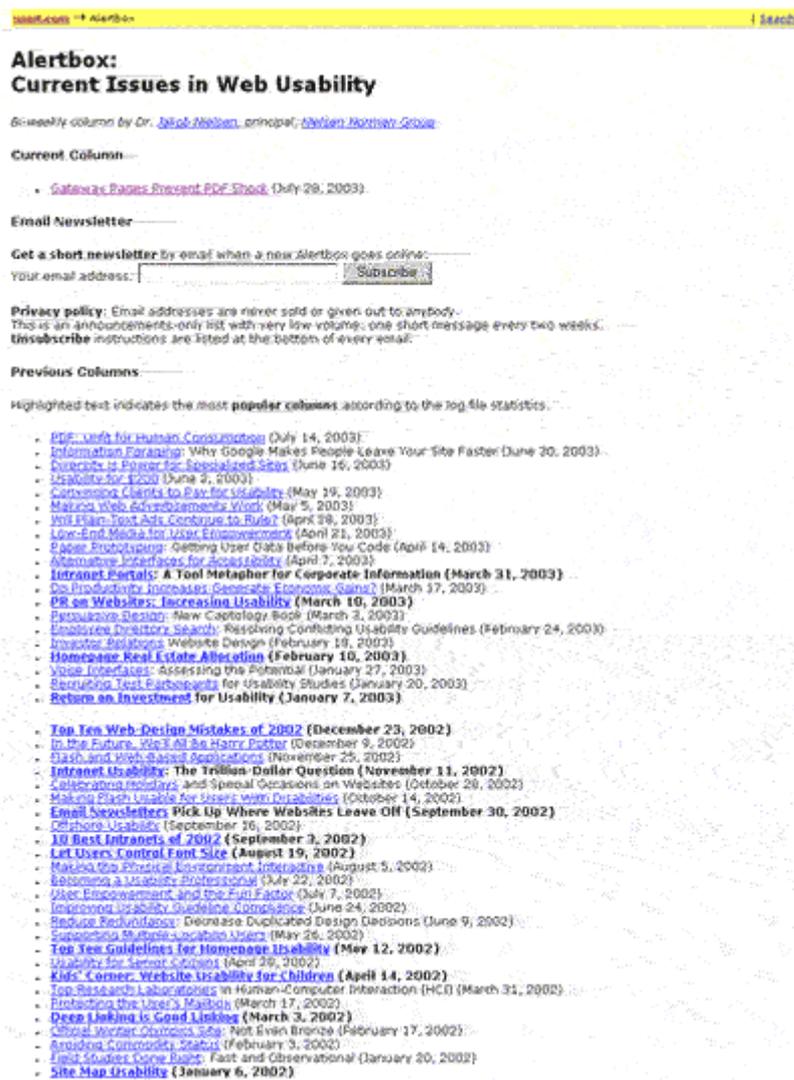
Guida Architettura del sito
 Prof. Aldo gorla
sitographics@infinito.it
<http://digilander.libero.it/sitographics>

Il principe dell'usabilità Jacob Nielsen ha linguaggio forbito e non manca mai di idee. C'è che lo detesta e chi lo impugna come un'arma spiegata contro ogni forma di creatività (spesso degli altri). Ma al di là dei contenuti, più o meno legittimi, l'organizzazione degli stessi poteva essere più rafforzata.

I famosi Alert-box sono tutti archiviati in un lunghissimo elenco (nella foto se ne vede uno screen capture a caratteri ridotti). La lista dei circa 230 titoli è mitigata dall'utilizzo sapiente di spazi e grassetti, ma non identifica macroaree di contenuto. Nel tempo si sono probabilmente formati dei filoni tematici, percorsi sviluppati in più interventi, iniziati e poi esauriti, su un dato argomento. Chi vuole ricostruire un percorso deve farlo a mano, spulciando il lungo indice (per esempio, l'argomento leggere e scrivere per il web compare in più articoli, ma dovremo essere noi a trovarli).

Più avanti invece vedremo come siti di grandi dimensioni, ma soprattutto siti con piccoli budget riescano a garantire un buon grado di ecologia dell'informazione.

Figura 5



I blog fanno venire il mal di testa

I blog, gli ormai famosi diari di bordo online compaiono su tutti i settimanali e quotidiani come il fenomeno del web. Di certo i blog aiutano chi li scrive, ma non aiutano chi li legge!

Non sono organizzati in nessun modo, sono un lungo serpentine cronologico di brani che i server gratuiti più diffusi (splinder, blogger) non permettono di archiviare diversamente.

Questo esempio (<http://spezzaincantesimi.splinder.com>) è stato scelto a caso partendo da un link sulla home page di splinder, impossibile rappresentarlo con uno screen shot data l'incredibile lunghezza. I blog sono diari personali, e ognuno nel suo diario fa ciò che vuole, ci mancherebbe. La cosa diventa più rischiosa se la tendenza da personale dovesse diventare pubblica! Allora l'intero web rischierebbe di far venire il mal di testa ...

L'"Information Brokering"

Più volte si è detto che l'efficacia di un sito sta nella possibilità di arrivare all'informazione desiderata. Il parametro un tempo era il numero dei click necessari: con tre click si doveva arrivare all'informazione. Questa definizione parte dal presupposto che l'utente abbia sempre in mente che cosa sta cercando, e inoltre che non vi siano possibili analogie al contenuto cercato.

Nell'accezione di cui si parla in questa guida, la **reperibilità dell'informazione** si misura con i requisiti fondamentali espressi da Tasso e Romero nella loro Fabbrica delle Informazioni (Personalizzazione dei contenuti web, Franco Angeli 2002).

L'architettura delle informazioni consiste quindi nel:

1. Fornire all'utente solo contenuti rilevanti.
2. Se i contenuti rilevanti sono troppi filtrare i più rilevanti.
3. Fornire un contenuto non appena è stato prodotto ed è a disposizione.
4. Fornire ciascun contenuto solo agli utenti interessati.
5. Fornire ciascun contenuto nel supporto tecnologico in cui si trova in quel momento l'utente finale.

Questi principi riuniscono insieme i principi di usabilità, accessibilità e information architecture.

I concetti chiave per arrivare a un vero e proprio sistema intelligente che argina gli effetti dell'overload informativo e allo stesso tempo rende un vero e proprio sistema di information brokering.

I principi dell'information brokering

Analisi e identificazione dei contenuti

I contenuti devono poter essere identificati mediante alcuni elementi fondamentali:

1. La classificazione generale a cui appartengono
2. Il titolo che li caratterizza
3. Una breve descrizione
4. Le parole chiave che li contrassegnano

Un altro significativo esempio: La pagina "Diritto allo studio" esiste nel portale dell'Università di Bologna, solo che facendo una ricerca testuale non emerge nessun risultato.

Non sempre è possibile (o opportuno per la navigazione) stabilire a priori tutte le possibili ramificazioni. Ecco perché è sempre più importante determinare un'organizzazione dal basso all'alto, che parta quindi dall'unità finale di contenuto e lo marchi con parole chiave (4).

In altre parole: ogni unità di contenuto, notizia o prodotto che sia, deve avere molteplici chiavi di relazione con il resto del sito.

Filtraggio dei contenuti

Dato un insieme di argomenti di interesse deve essere possibile valutare quanto un contenuto si avvicini ad essi, in modo da filtrare i contenuti considerati non rilevanti.

Ordinamento dei contenuti

Dato un argomento di interesse e un contenuto, deve essere possibile valutare il livello di rilevanza del contenuto relativamente all'argomento di interesse.

Raccomandazione di contenuti

Dato un insieme di argomenti di interesse deve essere possibile identificare e suggerire altri contenuti che appaiono potenzialmente interessanti.

Riformulazione di contenuti

Dato un contenuto e un insieme di caratteristiche dell'utente deve essere possibile modificare il contenuto al fine di adattarne la presentazione all'utente. Deve essere possibile offrire all'utente diverse modalità di fruizione e di accesso ai contenuti soprattutto in virtù del fatto che vi sono diversi supporti tecnologici e diverse abilità nel pubblico di riferimento; si rende sempre più necessario quindi separare la presentazione dal contenuto. Tutto ciò ha importanza soprattutto all'interno del tema dell'accessibilità.

Alerting dei contenuti

Dati contenuti particolarmente rilevanti deve essere possibile avvisare tempestivamente l'utente dell'esistenza di tali contenuti. (avviso degli utenti con inserti "what's new" e mailing)

Gli schemi percettivi e gli schemi esatti

Abbiamo visto che, quando i contenuti si accumulano, necessariamente occorre seguire una schematizzazione. Dagli esempi precedenti, si vede che una sola forma di schematizzazione non è sufficiente, occorre combinarla con altri tipi di schemi. Più gli schemi sono intersecati tra loro, più aumenta il grado di reperibilità di un'informazione. Ecco quindi una lista di tutti i possibili schemi di classificazione delle informazioni.

Schemi Percettivi (interfacce)

Le home page dei comuni portali dispongono i contenuti centrali in una forma non ordinata con criteri logici, ma visivi. Si tratta di schemi ricostruiti unicamente in base alla percezione, non hanno bisogno di troppe spiegazioni, hanno la caratteristica di avere una limitata serie di possibilità. Secondo alcune regole di psicologia delle interfacce, il numero di oggetti sparsi che la nostra percezione interpreta ancora come "ordine" è circa cinque.

Schemi Esatti (indicizzazioni)

Gli schemi organizzativi esatti dividono le informazioni in settori precisi e univoci la cui consultazione prevede una sola clausola: il visitatore deve avere ben chiaro in mente che cosa sta cercando.

- **Schemi esatti di tipo alfabetico**

Sono esatti gli schemi alfabetici, come gli elenchi di nomi propri, di paesi, di città. Lo schema alfabetico è adatto a queste tipologie di ricerca, mentre negli altri casi è solo di supporto.

Figura 6



- **Schemi esatti cronologici**

Particolarmente adatti agli uffici stampa online, sono in sostanza inutilizzata in siti con cataloghi di prodotti. Conoscere la data di inserimento di un prodotto non è così importante come conoscere la disponibilità in un determinato momento. In questo caso, l'unica indicizzazione cronologica è limitata alla link "novità" e "offerte" queste sì, temporalmente limitate.

- **Schemi esatti geografici**

Alcune tematiche politiche o socio/economiche, i viaggi, le reti di vendita fanno uso di schemi di navigazione per paese o regione: nei siti sul turismo e sul meteo sono indispensabili. Rispetto a un semplice schema alfabetico, recano un'informazione aggiuntiva: il territorio di appartenenza. E' importante aver cura di non lasciare solo alle mappe grafiche il compito di indicare i percorsi, il testo non può mai essere omesso.

Figura 7



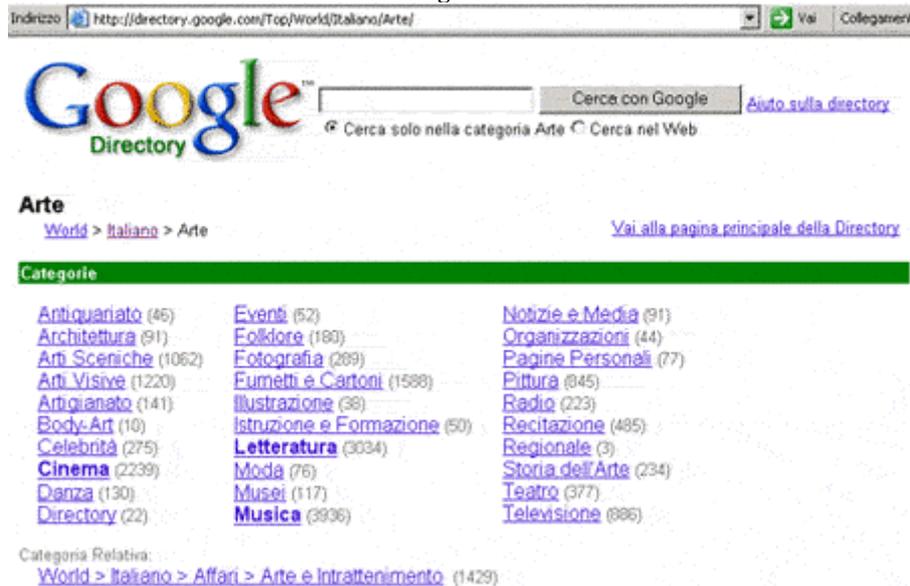
Gli schemi ambigui

Sono i veri schemi, ottenuti raggruppando e separando contenuti in categorie e sottocategorie, non sempre oggettive. Gli schemi organizzativi per titolo e autore sono esatti, non possono, se non in rari casi, dare adito a risultati inaspettati. Ma allo stesso tempo gli schemi ambigui come il codice Dewey, o la ricerca per argomento, sono indispensabili per gli utenti che stanno cercando di raccogliere più materiale possibile su una data materia di studio. Il motivo della fortuna degli schemi ambigui è molto semplice: non sempre sappiamo di preciso che cosa stiamo cercando. In alcuni casi il desiderio della ricerca è così vago da non saperlo esprimere. Inoltre, la ricerca in un sistema ben connesso è già apprendimento.

- **Schemi ambigui per argomento**

Per essere efficaci, le organizzazioni per argomento devono riferirsi a un contesto circoscritto. Pensare di catalogare l'intero scibile umano in directory, moltiplicare all'infinito le suddivisioni in sottocategorie non può essere un obiettivo realistico di un architetto dell'informazione. Cercare di circoscrivere un'area e suddividerla per argomenti e relazioni, invece, sì. Es.: (Google arte).

Figura 8



- **Schemi ambigui per utenza**

Particolarmente utilizzata nei siti di grandi aziende, la suddivisione per utenza ("Aziende", "Famiglie", "Rivenditori"...), presuppone una serie di informazioni mirate, servizi e impostazioni già esistenti di cui il web è il naturale riflesso. Nella pubblica amministrazione, alla tipologia di utenza si associa anche la navigazione per "Eventi della vita" (Nascita, matrimonio, trasloco, etc...): uno schema ambiguo, che ha il pregio di ricostruire tutti i passaggi necessari da compiere in un data situazione (Es: sito Ibm, siti della P.A.). La principale criticità di una schematizzazione per utenza è la capacità del visitatore di riconoscersi in un gruppo. Difficoltà che potrebbe riassumersi in questa domanda: "Ma io so chi sono io?" Esempio di suddivisione per utenza

Figura 9



- **Schemi ambigui per metafore**

Molto utilizzate nei sistemi operativi: desktop, cartella, cestino aiutano a comprendere intuitivamente contenuti e funzioni. Hanno avuto il loro momento di gloria nei primi siti internet, anche italiani, in cui le cosiddette "reti civiche" organizzavano la navigazione del sito cittadino come una grande mappa di una città in stile SimCity (Ospedale, teatro, scuola) . Nel momento in cui le relazioni si fanno complesse e la quantità di contenuti aumenta, la mappa cade in disuso, per lasciare il posto alle suddivisioni per argomenti e per utenza. unica metafora sopravvissuta: il carrello dei siti di E-commerce.

Gli schemi ambigui sono appunto imprecisi, non sempre oggettivi, ma questa ambiguità è la loro ricchezza.

L'organizzazione in categorie non è mai certa, come testimonia l'esempio di un gruppo di lavoro: trenta persone, divise in tre gruppi da dieci, hanno separatamente curato l'organizzazione in categorie di una stessa lista di contenuti, ottenendo tre risultati diversi.

un esempio di schema ambiguo. Come gli allievi del master Web Content Management, hanno organizzato le categorie della Regione Emilia-Romagna" utilizzando il metodo del card sorting:

Gruppo A	Gruppo B	Gruppo C
Istruzione Scuola università Formazione	Piano economico tributi Bilancio	Regione Organizzazione Statistica

Attività produttive	Statistica	Bilancio
Agricoltura	Ambiente	Territorio
Turismo	urbanistica	Piano territoriale
Artigianato	Piano territoriale	Edilizia
Industria	Cartografia	urbanistica
Commercio	Geologia	Sicurezza urbana
Politiche sociali	Protezione civile	Cartografia
Politiche sociali	Sicurezza urbana	Ambiente
Immigrazione	trasporti	Geologia
Emigrazione	Formazione/lavoro	Protezione civile
Pari opportunità	università	Istruzione
Giovani	Scuola	Giovani
urbanistica	Formazione	Scuola
urbanistica	Giovani	università
Edilizia	Lavoro	Concorsi
Piano territoriale	Servizi sociosanitari	Formazione e lavoro
Cartografia	Sanità	Tempo libero
Geologia	Pari opportunità	Cultura
Sicurezza	Politiche sociali	Sport
Protezione civile	Immigrazione	Turismo
Sicurezza	Emigrazione	Politiche sociali
Politiche locali	Comunicazione	Pari opportunità
Federalismo	Informatica	Sanità
Autonomie locali	Telematica	Politiche internazionali
Politiche estere	Tempo libero	Relaz. internazionali
Programmi uE	Sport	Programmi uE
Relaz. internazionali	Cultura	Immigrazione
Organizzazione	Turismo	Emigrazione
Organizzazione	Risorse economiche	Autonomie locali
tributi	Commercio	Federalismo
Bilancio	Industria	Politiche economiche
Statistica	Edilizia	Industria
Cultura e sport	Agricoltura	Commercio
Cultura	Artigianato	Artigianato
Sport	Politiche interne e	Agricoltura
Comunicazione e n. Tecnologie	Politiche esterne	trasporti
Informatica	Autonomie locali	Comunicazione
Telematica	Federalismo	Informatica
Lavoro	Programmi uE	Telematica
trasporti	Relaz. internazionali	
Ambiente		
Sanità		

Come classificare: card sorting cartaceo

Come si ottengono gli schemi di classificazione ambigui? Uno dei metodi più seguiti, [suggerito](#) a suo tempo da Jacob Nielsen, è quello del Card Sorting. Si tratta di una elaborazione in categorie di un insieme di carte, operata da gruppi di esperti del settore, in cui l'architetto dell'informazione svolge il compito di sintetizzare e mediare tra le risposte ottenute.

Un card sorting può essere svolto attraverso una serie di foglietti o carte (metodo cartaceo) o attraverso software (metodo digitale).

Figura 10: Un esempio di card sorting



- Creare carte in cui vengono riportati gli argomenti che riguarderanno il sito (eventualmente si possono concordare con gli esperti i termini esatti da utilizzare nelle carte).
- Chiedere agli esperti di raggruppare le carte in gruppi logici.
- Analizzare i gruppi che gli esperti creano e determinare l'organizzazione ottimale della struttura nel sito.

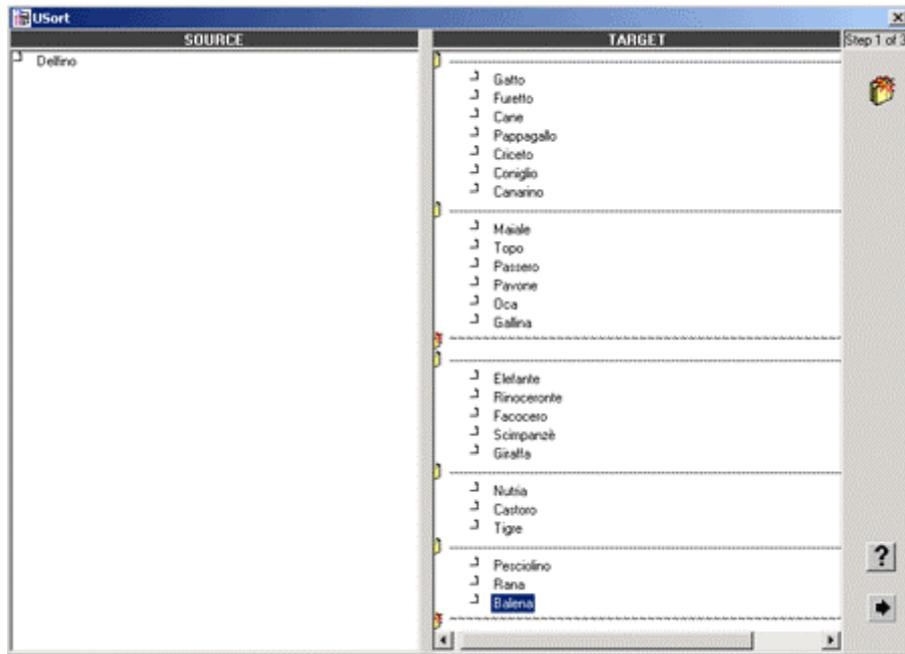
Se i contenuti del sito sono generici, invece che a esperti si può ricorrere al parere di utenti rappresentativi che si suppone appartengano al pubblico di destinazione.

Come classificare: card sorting digitale

IBM ha sviluppato EZSort, un software per la gestione di card sorting digitali. È un programma che consente di scrivere una lista di nomi (in questo esempio nomi di animali) e di organizzarli per gruppi in maniera molto semplice.

trascinando i nomi che compaiono nella colonna di sinistra in aree della colonna di destra, è possibile operare dei raggruppamenti

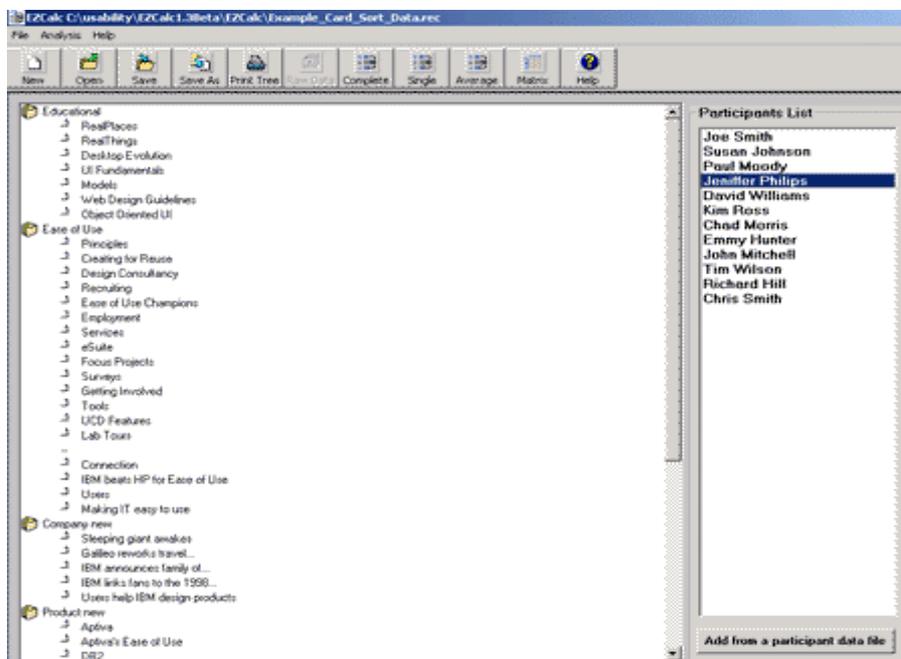
Figura 11



Una volta che i componenti del gruppo di lavoro hanno fatto le loro scelte, si utilizza l'applicazione Ez Calc per rappresentare i raggruppamenti operati.

Nell'esempio: le scelte di Jennifer Philips

Figura 12



La media delle scelte

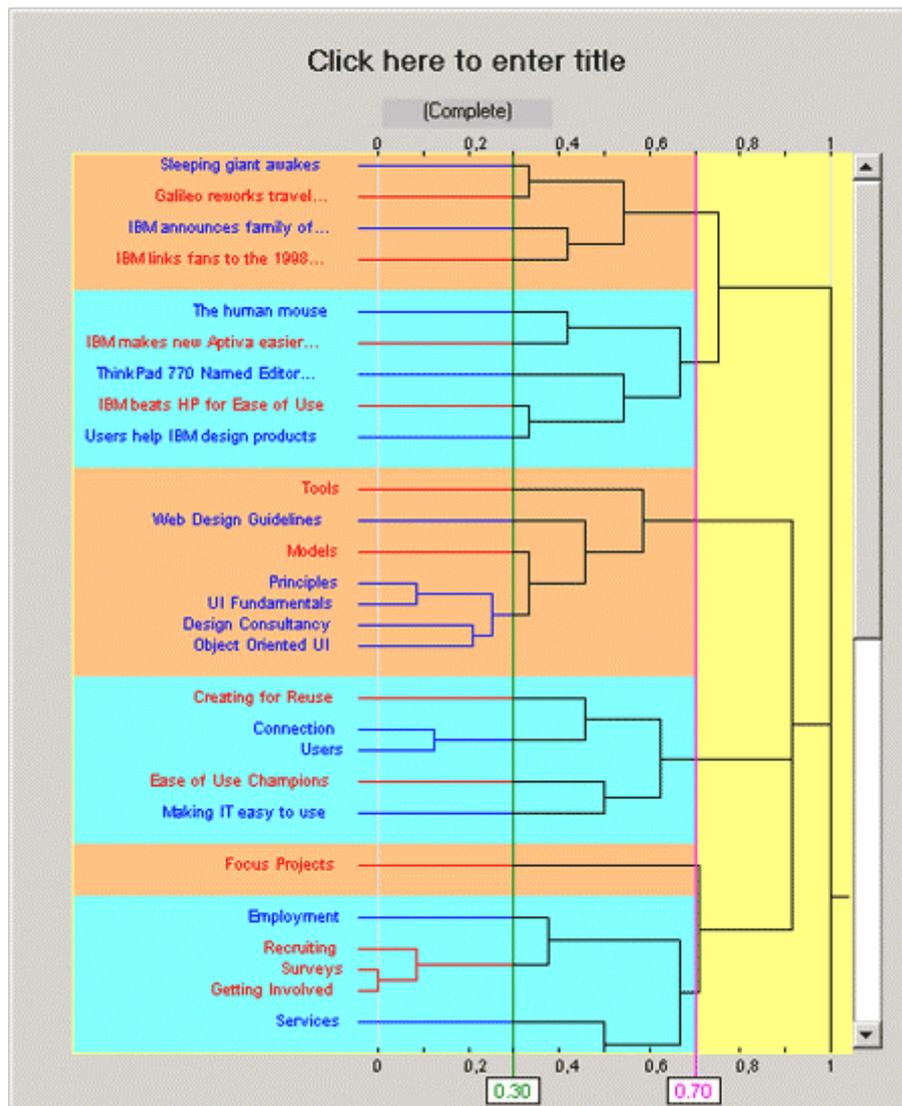
Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

Figura 13



Dalle liste di raggruppamenti ottenute da più partecipanti è possibile ottenere alberi di categorie. Linee di diverso colore indicano che è stato usato un algoritmo di calcolo differente.

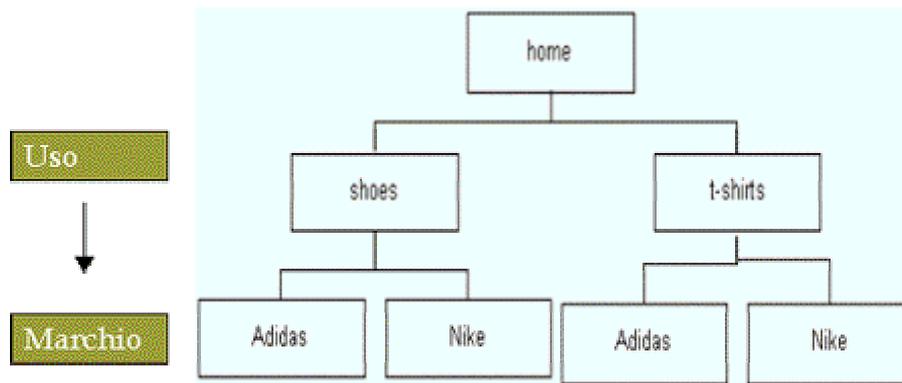
Dopo il Card Sorting: il dilemma delle categorie

Anche quando si sono individuate le categorie, restano aperti nuovi interrogativi, soprattutto quando si intravedono molteplici possibilità di classificazione. A quale dare la precedenza?

Un esempio classico è quello dei siti di moda in cui è relativamente semplice suddividere i prodotti per categoria di appartenenza (t-shirt, gonne, scarpe, etc). Ma un altro criterio di ingresso importantissimo, nella moda, è quello per stilista.

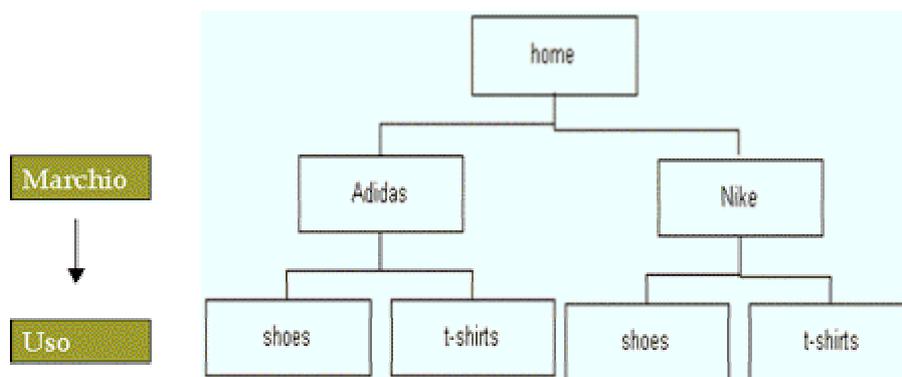
Dall'uso al marchio?

Figura 14



Dal marchio all'uso?

Figura 15



Yoox, uno dei più famosi siti di e-commerce offre la scelta per brand e la scelta per categorie:

Figura 16



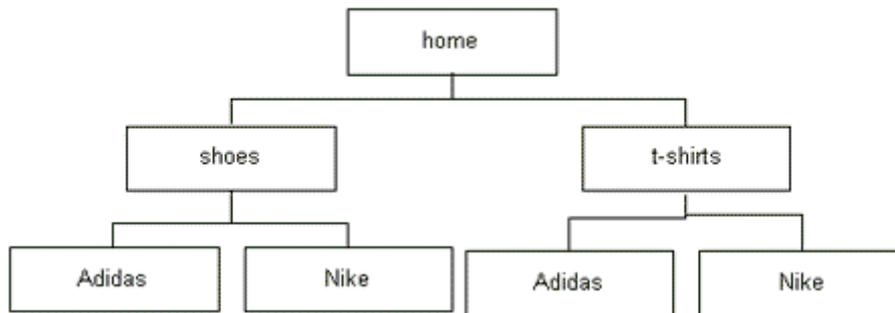
In questo caso è necessaria una mediazione, sarà possibile individuare un doppio ingresso per arrivare agli stessi prodotti:

La gerarchia ad albero: un approccio dall'alto al basso

Si pensa a una gerarchia ad albero quando si pensa al sito già pronto da navigare.

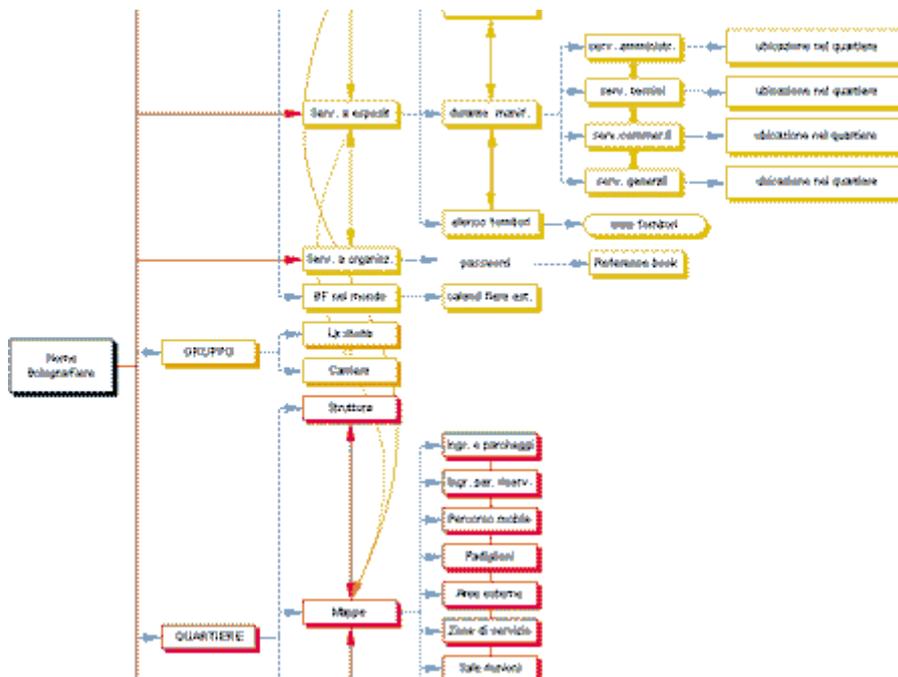
Le pagine e gli accessi sono gerarchici, categorie e sottocategorie sono definite in base a diagrammi di flusso che ricalcano visivamente la mappa del sito.

Figura 17



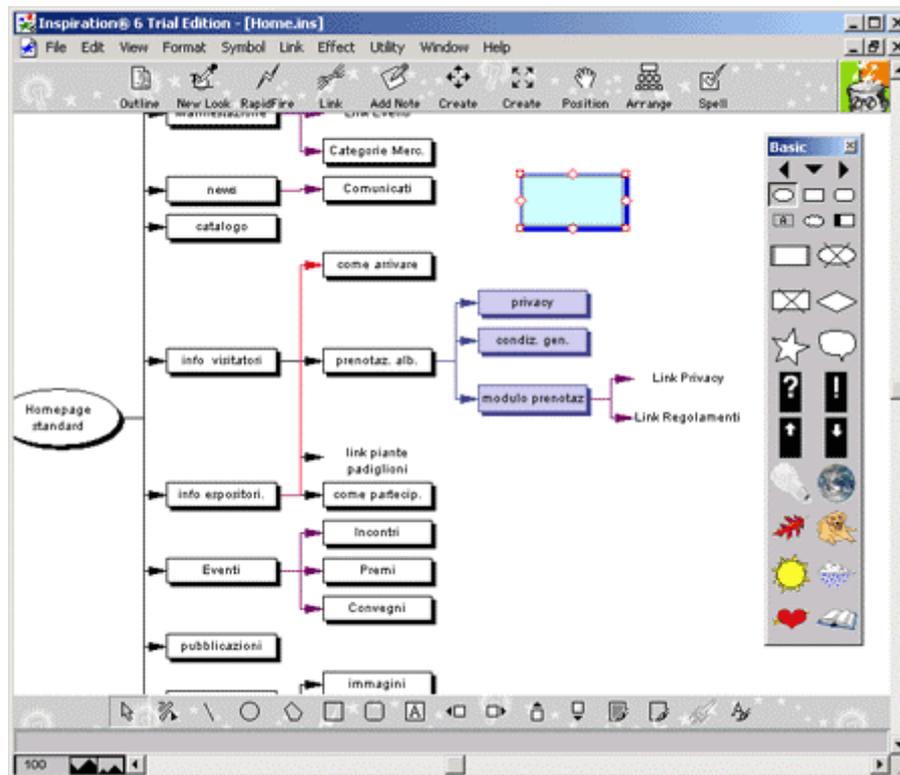
La definizione degli schemi (ambigui) di classificazione ci porta a costruire l'albero di navigazione, che è in ultima analisi, la definizione del sito in tutta la sua ampiezza.

Figura 18



Un software utile per la creazione dell'albero di navigazione è Inspiration.

Figura 19



Architettura dei task e timone dell'homepage

Il software Inspiration, citato nella lezione precedente, può essere utilizzato anche per gestire i task **di un sito**, e non solo la mappa di navigazione. Che cosa si intende per task?

I task sono tutte quelle azioni che l'utente deve compiere per arrivare a un determinato risultato, non necessariamente si tratta di pagine, ma di eventi che si susseguono, di "se, allora.." che vanno predefiniti con estrema cura.

Nel caso riportato nella figura 20 qui in basso vediamo ricostruito il flusso di una finestra di log-in e registrazione ad aree riservate del portale. Sembrano percorsi scontati: ma non è sempre così.

Questo flusso indica chiaramente che l'utente deve sempre poter vedere il menu e il contenuto della pagina indice.

Se è già loggato, cliccando su un link di suo interesse il sistema lo manderà direttamente alla pagina desiderata. Altrimenti gli proporrà di autenticarsi o di registrarsi ex novo. Il tutto senza mai far sparire il contenuto della pagina di partenza che è il vero motivo per cui l'utente decide di registrarsi e che non deve mai essere posta in secondo piano.

Lo schema che segue indica che tutto deve avvenire in una finestrella che si apre al momento del click su un link, una popup accessibile e non invasiva, che lasci sempre intravedere il contenuto della pagina, senza costituire un vicolo cieco, e che infine, non consenta di proseguire in assenza della dovuta registrazione. Questo tipo di flusso non a vicolo cieco ha favorito l'aumento delle registrazioni nel portale.

Figura 21



Il Content Management System e i metadati

L'architetto delle costruzioni progetta le sue realizzazioni su solidi pilastri di cemento armato. L'architetto delle informazioni affida le strutture portanti del suo sito ai **metadati**. Non si tratta di accumulare un altro insieme di dati da giustapporre alla somma delle informazioni. L'obiettivo è **concatenare le informazioni** e i dati che formano ogni scheda finale di contenuto. Le schede finali di contenuto rientrano in genere in due tipologie principali: prodotto (per esempio, la scheda di un abito in un sito di e-commerce), o notizia (per esempio, un articolo in un portale informativo, verticale o generalista).

Il luogo in cui materialmente si interviene sull'unità finale di contenuto è il **content management system (CMS)**: un sistema per la gestione di contenuti informativi da pubblicare su siti e portali web. Se andiamo a leggere le guide e i tanti articoli americani sull'architettura delle informazioni e il content management, leggeremo che una, l'architettura, ha una visione immediata delle informazioni (sincronica), mentre l'altro, il content management le vede inserite in un flusso cronologico (diacronica). A parte l'eleganza delle distinzioni, io sono personalmente dell'idea che questi due aspetti non possano essere separati, né operativamente né professionalmente.

Il sistema di content management è la **realizzazione pratica dei principi di information architecture**, e allo stesso tempo, ne è l'estensione. Realizzarlo richiede delle grandi capacità di programmazione, ma il ruolo dell'Information Architect non è scrivere del codice bensì seguire la progettazione di un cms, o il processo di personalizzazione se si tratta di un prodotto già esistente sul mercato. Quando un buon sistema di content management è al lavoro con un gruppo di validi redattori, il risultato finisce per stupire anche chi ha pensato il progetto iniziale.

Vedere il contenuto "al lavoro", mentre si interseca e si incontra in rivoli di connessioni e associazioni, che inaspettatamente emergono da ogni parte, supera sempre le attese iniziali. E senza un buon lavoro (umano) di redazione ogni architettura delle informazioni è debole.

Vediamo cosa succede nella pratica: come un'architettura delle informazioni si realizza in un sistema di content management!

Figura 22

Esempio visivo di scheda di contenuto all'interno di un sistema di Content Management.

Categoria padre	Economia dell'edilizia <input type="checkbox"/> La categoria principale
Categoria	Nuovi mercati e internazionalizzazione <input type="checkbox"/> la sottocategoria
Visibilità nei menu	Notizia visibile nei percorsi di navigazione? SI <input checked="" type="checkbox"/> NO <input type="checkbox"/>
Data di scrittura	29 / 09 / 2003 La data in cui la notizia è stata scritta
Data di pubblicazione	06 / 10 / 1972 La data in cui la notizia deve essere pubblicata (se diversa dalla data di scrittura)
Notizia temporanea Si/No	E' un contenuto: Temporaneo <input type="checkbox"/> Permanente <input checked="" type="checkbox"/>
Scadenza notizia gg/mm/anno	12 / 200 / 2003 Se la notizia è temporanea indicarne la scadenza
Lingua	Italiano <input type="checkbox"/> Se il sito è localizzato in più lingue anche questa informazione va aggiunta
Titolo	La Cina è a sorpresa il mercato di riferimento. Titolo della notizia, è meglio non superare i cento caratteri. Il titolo della notizia sarà anche il titolo della pagina .
Sottotitolo (descrizione breve)	La Cina tra mille contraddizioni, si prefigura come uno dei maggiori paesi per l'investimento Lo stillo o abstract contiene una descrizione breve della notizia, vi sono già contenute alcune parole chiave di riferimento per trovare la notizia con i motori di ricerca nel sito
Testo integrale	Chengdu, provincia del Sichuan, il vicespagnolo della Commissione statale per la pianificazione dello sviluppo e vicedirettore dell'Ufficio per lo sviluppo della Cina dell'ovest sotto il Consiglio di Stato, Wang Chunzheng, ha annunciato un pacchetto di nuove politiche governative che hanno lo scopo di spronare lo sviluppo della Cina dell'ovest e di incoraggiare la cooperazione tra la Cina dell'ovest e il resto del mondo. Con queste politiche si vuole incrementare l'introduzione di capitali, migliorare l'ambiente d'investimento, allargare l'apertura all'esterno, attrarre persone di talento e sviluppare istruzione, scienza e tecnologia.
Responsabile della categoria	Paolo Degli Esposti Questo campo non è modificabile, sarà il referente che autorizza la pubblicazione di questa notizia
Firma dell'autore	Emanuela Gugnelli Firma della news: testo di max caratteri
Parole chiave in uscita	piscina, alluminio, legno, pietra
Parole chiave in entrata	piscina, alluminio, legno, pietra
Thesauro tematico	Esportazioni Estremo oriente Voci di thesauro
Pubblico di riferimento	Principianti Esperto A chi si rivolge questa notizia? Selezionare uno o più gruppi

Metadati di categoria

Ogni contenuto del sito viene inserito nel content management system attraverso una scheda, composta da numerosi campi: analizziamo i vari tipi di metadati che bisogna prevedere in un buon CMS.

Metadati di categorizzazione

Categoria padre

La categoria principale

Esempio: Economia dell'edilizia

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

Categoria	Una ulteriore suddivisione	Esempio: Nuovi mercati e Internazionalizzazione
Il contenuto è visibile nei percorsi di navigazione? (Si/No)	Utile per creare pagine di servizio (es: la parte seconda di un articolo molto lungo) non inserite nei menu di navigazione	

Figura 22°

Categoria padre	Economia dell'edilizia	La categoria principale
Categoria	Nuovi mercati e internazionalizzazione	la sottocategoria
Visibilità nei menu	Notizia visibile nei percorsi di navigazione? SI <input checked="" type="checkbox"/> NO <input type="checkbox"/>	

Questi elementi indicano in che punto della navigazione si trova il contenuto. In quale categoria e sottocategoria (volendo i livelli possono anche essere di più).

In questa parte si definisce la posizione logica del contenuto, la sua "residenza fissa".

Come abbiamo già avuto modo di notare, uno dei problemi è che, se la categoria "Nuovi mercati" all'interno della categoria padre "Economia" contiene più di trenta articoli, diventa estremamente difficile rintracciare un contenuto tra i tanti. Se poi mettiamo insieme notizie che hanno un'importanza nel tempo e notizie spot che passano velocemente, la confusione aumenta ancora. Per questo motivo si aggiungono altri tipi di metadati.

Metadati di tempo

Aggiungono al contenuto dati temporali come la data di pubblicazione e la data di scadenza. Se un contenuto giunge alla data di scadenza che l'editore ha indicato, allora scompare dagli alberi di navigazione per categorie, limitando così le possibilità di sovraccarico visivo.

I metadati Time Sensitive

Data di scrittura	La data in cui il contenuto è stato scritto e inserito nel CMS
Data di pubblicazione	La data in cui il contenuto deve essere pubblicato, se diversa dalla data di scrittura: opzione utile per preparare in anticipo i materiali.
Contenuto temporaneo (Si/No)	È una notizia temporanea o permanente?
Data di scadenza contenuto	Se la notizia è temporanea indicarne la scadenza: il CMS provvederà automaticamente a toglierla dal sito.

Figura 22b

Data di scrittura	29 / 09 / 2003	La data in cui la notizia è stata scritta
Data di pubblicazione	06 / 10 / 1972	La data in cui la notizia deve essere pubblicata (se diversa dalla data di scrittura)
Notizia temporanea Sì/No	E' un contenuto: Temporaneo <input type="checkbox"/> Permanente <input checked="" type="checkbox"/>	
Scadenza notizia gg./mm./anno	12 / 200 / 2003	Se la notizia è temporanea indicarne la scadenza

Un contenuto time sensitive può essere eliminato su più fronti:

1. dalla navigazione per categorie: indispensabile per temi che scadono come appuntamenti, convegni, seminari, ecc.;
2. dalla ricercabilità tramite motori di ricerca interni. È meglio far scomparire completamente informazioni di servizio scadute, del tipo "per la presentazione dell'attività 2001 appuntamento all'Hotel Plaza alle 11.00". Si possono invece lasciare tra i risultati di una ricerca testuale, se contengono informazioni aggiuntive a quelle di servizio (note sui relatori, piccoli abstract).

Inoltre, all'interno di una stessa categoria di argomenti/informazioni dovrebbe essere sempre possibile prevedere articoli a scorrimento veloce (a mo' di live news) e a scorrimento lento (dossier); anche visivamente questi contenuti devono essere elencati in maniera differente all'interno di una stessa categoria. Vediamo un esempio:

Metadati linguistici e redazionali

Lingua Se il sito ha versioni in più lingue, è indispensabile aggiungere anche questa informazione per ogni singolo contenuto.

Metadati redazionali

Titolo	Titolo del contenuto, è meglio non superare i cento caratteri. Il titolo della contenuto sarà anche il titolo della pagina.	Esempio: "Un nuovo Linux: rilasciato il kernel 2.6 "
Sottotitolo (descrizione breve)	Lo strillo o abstract contiene una descrizione breve del contenuto, vi sono già contenute alcune parole chiave di riferimento per trovare la pagina con i motori di ricerca del sito	Esempio: "Un nuovo Kernel per un nuovo Linux. Linus Torvalds ha annunciato oggi la versione definitiva del Kernel 2.6 che apporta novità e proietta Linux verso il futuro. Per le prime distribuzioni commerciali '2.6' ci vorranno mesi di attesa "
Testo integrale	Il testo integrale del contenuto o corpo	

Figura 22c

Lingua	Italiano <input type="button" value="v"/> Se il sito è localizzato in più lingue anche questa informazione va aggiunta
Titolo	La Cina è a sorpresa il mercato di riferimento. Titolo della notizia, è meglio non superare i cento caratteri. Il titolo della notizia sarà anche il titolo della pagina .
Sottotitolo (descrizione breve)	La Cina tra mille contraddizioni, si prefigura come uno dei maggiori paesi per l'investimento <input type="button" value="v"/> Lo stillo o abstract contiene una descrizione breve della notizia, vi sono già contenute alcune parole chiave di riferimento per trovare la notizia con i motori di ricerca nel sito
Testo integrale	Chengdu, provincia del Sichuan, il vicespagnabile della Commissione statale per la pianificazione dello sviluppo e vicedirettore dell'Ufficio per lo sviluppo della Cina dell'ovest sotto il Consiglio di Stato, Wang Chunzheng, ha annunciato un pacchetto di nuove politiche governative che hanno lo scopo di spronare lo sviluppo della Cina dell'ovest e di incoraggiare la cooperazione tra la Cina dell'ovest e il resto del mondo. Con queste politiche si vuole incrementare l'introduzione di capitali, migliorare l'ambiente d'investimento, allargare l'apertura all'esterno, attrarre persone di talento e sviluppare istruzione, scienza e tecnologia.

Titolo.

È una fonte inesauribile di parole chiave!

Noi inseriamo metadati anche senza saperlo. Nel momento in cui strutturiamo una unità di contenuto con un titolo e un abstract, per esempio, abbiamo già compiuto un lavoro di indicizzazione della nostra unità di contenuto.

In un sistema di ricerca il titolo è uno dei campi più importanti, è il campo che ha più peso nella determinazione dei risultati, vi andrebbe messa molta attenzione.

Evitare vocaboli generici come cosa, qualche, qualunque, alcuni; evitare anche metafore comuni: "casa del...", "calderone", che in realtà non sono pertinenti con il contenuto.

Si ai vocaboli specifici, anche tecnici, se sono in tema: per esempio termini come ristrutturazione edile, legno lamellare, sigillanti, vanno benissimo in un sito dedicato ai professionisti dell'edilizia.

Sottotitolo - descrizione breve

Contiene un'anticipazione riassuntiva dell'informazione, quella che costituisce l'accompagnamento del link (il cosiddetto "microcontenuto"). La funzione della descrizione breve è importantissima per la rintracciabilità dei testi. Questo campo contiene parole importanti, che durante lo scanning veloce della pagina aiutano il visitatore a capire di che cosa si sta parlando e che identificano il contenuto ai fini della ricerca con il motore interno al sito.

Testo integrale (il corpo del testo)

Contiene il corpo dell'articolo vero e proprio, il testo dell'unità di informazione. In genere è un campo poco valorizzato nei motori di ricerca interni al sito: proprio perché a causa della sua estensione può contenere termini che non hanno una pertinenza sostanziale e che rischierebbero, se considerati, di dare risultati fuorvianti in fase di ricerca.

Metadati di authoring e di target

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

Metadati di authoring

Si tratta di metadati riferiti all'autore e "pubblicatore" della notizia. Ovvero chi materialmente scrive e chi ha il permesso di pubblicazione. In siti complessi può avvenire che ogni categoria abbia il suo referente per la pubblicazione.

Responsabile della categoria	Ogni categoria ha un proprio responsabile, è la figura professionale che con la sua password privata può approvare la pubblicazione del contenuto. Se ritiene che il testo non sia pertinente alla categoria, può rifiutarlo e proporlo al responsabile di un'altra categoria.
Editor	È la figura professionale che inserisce il contenuto nel CMS. Può inserire sia testi scritti da lui sia testi scritti da una redazione.
Firma dell'autore	Il modo in cui compare nella pagina del sito il nome dell'autore del testo, con la funzione aziendale o l'indirizzo email.

Figura 22d

Responsabile della categoria	<input type="text" value="Paolo Degli Esposti"/> Questo campo non è modificabile, sarà il referente che autorizza la pubblicazione di questa notizia
Firma dell'autore	<input type="text" value="Emanuela Gugnelli"/> Firma della news: testo di max caratteri

Il pubblico di riferimento (target)

Sempre più siti, soprattutto quelli della pubblica amministrazione, offrono percorsi per target. Come già accennato, crediamo che non sia bene considerarlo come unico schema di navigazione, perché non sempre il nostro pubblico è completamente certo di rientrare in un determinato gruppo; ma può essere utile tenere questo tipo di schema come ingresso aggiuntivo, semplicemente contrassegnando ogni unità di contenuto come più adatta a un tipo di pubblico piuttosto che a un altro. Il database poi si preoccuperà di far "volare" dietro il link "Principianti" tutti i contenuti che sono stati contrassegnati come tali.

Figura 22e

Pubblico di riferimento	<input type="text" value="Principianti"/> <input type="text" value="Esperti"/> A chi si rivolge questa notizia? Selezionare uno o più gruppi
--------------------------------	---

Metadati

Il thesauro è un insieme di parole chiuso. Il redattore deve associare ogni contenuto ad una o più parole di questo elenco. Per la maggiore staticità, questo sistema associativo è spesso più adatto a unità di contenuto di tipo "prodotto".

Figura 22f

Parole chiave in uscita	<input type="text" value="piscina, alluminio, legno, pietra"/>
Parole chiave in entrata	<input type="text" value="piscina, alluminio, legno, pietra"/>
Thesaurus tematico	<input type="text" value="Esportazioni"/> <input type="text" value="Estremo oriente"/>

Voci di thesaurus

Il tipo più semplice di thesaurus è composto da un elenco di parole in ordine alfabetico. Esempio di thesaurus semplice che è gestito dal sistema di contenuti del sito Edilio.it:

Figura 23

	A	B	C	D	E	F
1	Thesaurus dell'edilizia					
2	Abbaio					
3	Abitazione					
4	Acciaio					
5	Acciaio inossidabile					
6	Acciaio: corrosione					
7	Accumulatore termico					
8	Acqua					
9	Acque reflue: depurazione					
10	Acque reflue: smaltimento					
11	Acquedotto					
12	Acustica					
13	Additivo					
14	Adesivo					
15	Aerofotogrammetria					
16	Aeroporto					
17	Albergo					
18	Allestimento					
19	Alloggio					
20	Alloggio: arredamento					
21	Alloggio: ristrutturazione					
22	Alluminio					
23	Alluminio: manutenzione					
24	Ambiente					
25	Ambiente: monitoraggio					

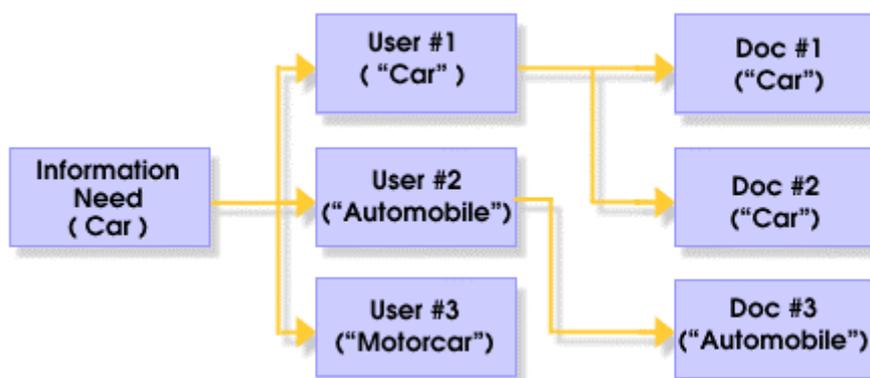
Esempio di Thesaurus semplice

Thesauri complessi

La realtà chi progetta e realizza siti costringe l'architetto delle informazioni a fare i conti con una maggiore complessità: molte informazioni rimangono inesplorate perché i termini di ricerca utilizzati dall'utente non coincidono con quelli inseriti nel documento da autori e organizzatori .

Figura 24

tratto da da http://www.webreview.com/1998/10_30/developers/10_30_98_1.shtml



Il thesaurus complesso è una collezione di nomi riferiti a una particolare disciplina e legati tra loro da una serie di relazioni semantiche. Ogni termine di un thesaurus viene posto al centro di una rete, in cui si distinguono tre tipi di relazione.

1. **Relazione di equivalenza**

a sinonimia nel senso classico del termine, es. "Onde" e "Flutti", "Aperitivo" e "Happy Hour". L'equivalenza è una relazione più forte del sinonimo, indica un rapporto di totale intercambiabilità tra due termini. Sebbene i linguisti concordino nell'affermare che non esistono due termini perfettamente intercambiabili, l'equivalenza, oltre a far coincidere il significato, vede minimizzate le differenze di stile e registro fra due termini.

2. **Relazione gerarchica**

: la classificazione dei termini in categorie e sottocategorie, ad esempio "Felino" e "Gatto" "Vino" e "Sangiovese". Tipi di gerarchia:

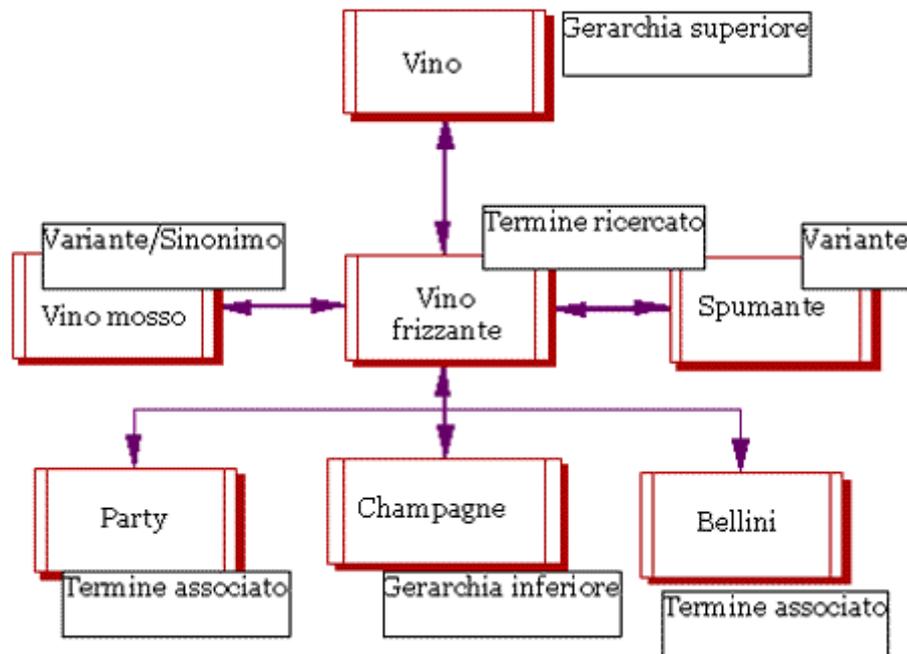
- o Gerarchia generica: la classica relazione classe-specie della biologia. La specie "gatto" eredita tutte le qualità della classe "felino".
- o Parte-Tutto: Il dito è parte della mano.
- o Esemplificativa: mari, Mar mediterraneo.

3. **Relazione associativa**

: offre connessioni ritenute significative e diverse da gerarchia e equivalenza. Per esempio: tra "Prosecco" e "Happy hour", la relazione associativa è detta anche residuale, ed ha un carattere fortemente soggettivo o legato alle abitudini sociali.

Le relazioni associative sono molto utilizzate nei siti di e-commerce: permettono di abbinare a un oggetto scelto, altri oggetti correlati. Chi acquista un vino raro in rete potrebbe avere interesse ad acquistare anche un formaggio altrettanto raro da abbinarvi. (Nell'immagine, un esempio liberamente tratto da Architettura dell'Informazione per il Word Wide Web)

Figura 25: tre tipi di relazione



I sistemi di navigazione

Un buon architetto delle informazioni non può fare a meno di conoscere tutti gli standard della navigazione, standard che di fatto costituiscono gli strumenti operativi del dialogo con il visitatore.

Come abbiamo visto in precedenza, lo scopo della navigazione è:

- presentare ai visitatori, con il percorso più user friendly, la classificazione degli elementi del sito, così che possano trovare velocemente il contenuto che cercano;
- assicurare che i visitatori sappiano sempre in che parte del sito si trovano;
- permettere ai visitatori di muoversi velocemente e in modo logico attraverso il sito web;
- dare ai visitatori i riferimenti appropriati del contenuto che stanno leggendo;
- fornire esempi di contenuti che l'organizzazione vuole promuovere.

Vediamo alcuni esempi di sistemi di navigazione:

1. Navigazione Centrale

: quella che utilizza la parte centrale del corpo dell'informazione: il contenuto della pagina coincide in gran parte con i link di navigazione. Un esempio è il sito yahoo.it, con le sue directory "Arte e Cultura", "Economia", "Società", etc... Nella parte centrale della pagina risiede spesso l'indice del sito, organizzato per categorie.

2. Navigazione Globale

: contiene il link alle pagine che devono essere accessibili da ogni parte del sito (Home, contatti, etc). La navigazione globale dovrebbe anche essere posizionata al di sopra e al di sotto di ogni pagina. Nel raggruppamenti per schemi e categorie, la navigazione contiene il primo livello di categorizzazione del sito, le macro categorie che contengono tutte le altre.

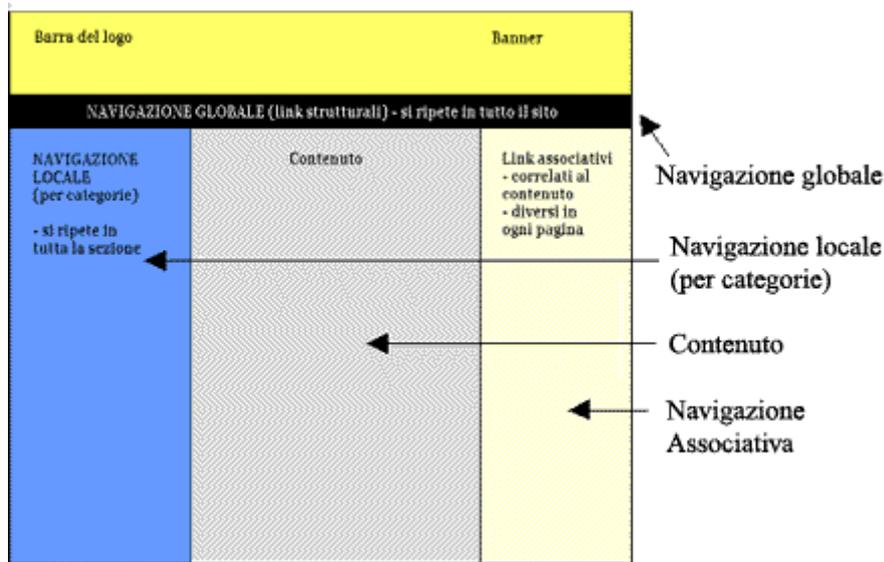
3. Navigazione caratteristica

: si usa questo tipo di navigazione per enfatizzare alcune attrattive di contenuto del sito web. Per esempio: "La storia del giorno", oppure un prodotto appena uscito, etc. Si tratta di prendere un prodotto o una notizia dall'albero di navigazione e, per motivi promozionali, renderlo visibile a un livello superiore.

4. **Navigazione locale:**
si utilizza in genere sul lato sinistro del documento. Riprende l'indice delle categorie collegate al documento che si sta consultando.
5. **Storia della navigazione:**
contiene esattamente la ricostruzione del percorso effettuato per giungere alla pagina in cui ci si trova, ed è fortemente correlata al sistema generale di classificazione. Dovrebbe sempre iniziare con il link della Home e finire con la sezione in cui ci si trova. (Per esempio, Home/Prodotti/Computer/Portatili). Nielsen ha dato a questa tecnica il fiabesco nome di "bread crumbs", briciole di pane.
6. **Navigazione associativa:**
è un menu contestuale ad un singolo documento che cita altre pagine appartenenti a categorie diverse. E' la forma visibile dell'associazione tramite metadati già ampiamente trattata in questa guida.
7. **Url navigation:**
la url è l'indirizzo di un documento su Internet, e consiste nella comunicazione di un protocollo seguito da due punti e due barre. Questa navigazione ci dice l'indirizzo preciso della pagina in cui ci troviamo. Alcuni utenti esperti la utilizzano per muoversi a ritroso nelle cartelle del sito, e inoltre è una caratteristica utile per segnalare la pagina ad altre persone, ad esempio via email.

Lo schema seguente mostra come disporre i vari tipi di navigazione nella pagina, secondo la prassi consolidata nel web.

Figura 26



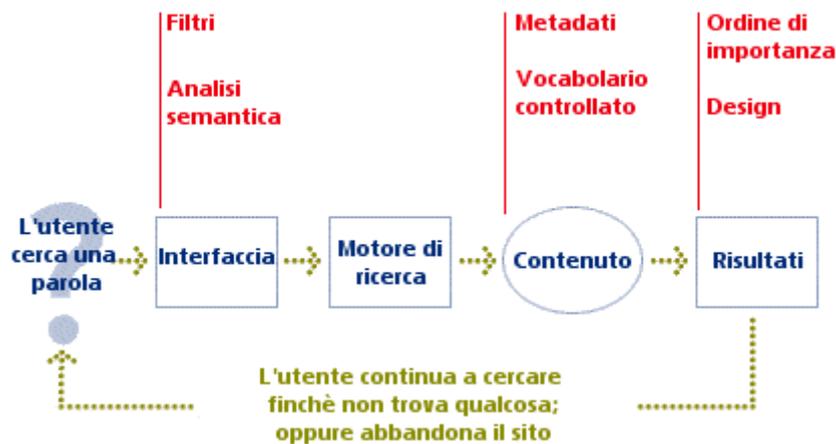
Analisi dei termini della ricerca

L'obiettivo da cui siamo partiti è quello di far sì che l'utente trovi sempre nel sito l'informazione che sta cercando. Nella fase di progettazione, quindi, dobbiamo dedicare la nostra attenzione anche alle funzionalità del motore di ricerca interno.

L'immagine che segue (ispirata a quella pubblicata nel sito

<http://semanticstudios.com/publications/semantics/000004.php> e riportata nel volume Architettura dell'informazione per il World Wide Web) sintetizza gli aspetti dell'AI che intervengono durante le fasi di ricerca:

Figura 27



Analizzare i termini di ricerca

L'utente scrive una o più parole nel campo di ricerca. Questa azione, in motori semplici, porta subito il sistema a scandagliare le unità di contenuto in cerca di possibili match, con risultati che possono essere deludenti, soprattutto in caso di ricerca di parole molto generiche. Per fornire un servizio efficace, dobbiamo usare innanzitutto un sistema complesso di analisi linguistica. Ne esistono di due tipi: semantici e comparativi.

L'analisi semantica dei termini si basa sui significati delle parole e sulle regole della logica. Dopo che la parola chiave è stata digitata e immessa, il sistema chiede all'utente di specificare meglio la sua ricerca, offrendogli alcune scelte: è come se facessimo consultare all'utente un dizionario dei sinonimi.

Un esempio sul sito TonyStone.com, un ricchissimo database di immagini.

Se ricerchiamo una parola semplice come "house", il motore di ricerca ci chiede di specificare se intendiamo "casa" come "edificio" oppure come "abitazione": il risultato sarà più vicino alle nostre aspettative.

Figura 28

Questi tipi di filtri di ricerca (query builders) presuppongono l'esistenza di software predefiniti, che analizzano il contenuto delle query per restituire risultati più efficaci.

L'altro tipo di analisi, quello comparativo, è visibile nella vita internetiana di tutti i giorni, infatti è utilizzato dal benemerito Google. Il motore di ricerca confronta tra loro le varianti ricercate simili, per ricondurle al termine maggiormente cercato. Se l'utente digita un nome di persona sbagliando una lettera, il sistema ripropone quello più richiesto, che probabilmente è anche quello esatto. Questa analisi, utilissima per correggere errori di ortografia, utilizza il principio del filtraggio collaborativo (che vedremo meglio più avanti).

Facciamo un esempio: immaginiamo di trasformarci in un adolescente con scarsa conoscenza dell'inglese e un grande amore per una stellina del pop... un po' frettolosamente, ricercheremo informazioni su "britni spirs". Google, con atteggiamento materno, ci indirizza subito al nome esatto: "Forse cercavi "Britney Spears".

Figura 29



Più interessante il caso in cui l'errore di digitazione sia meno macroscopico: ricercando "Brittany Spears", non si ha più la correzione, ma un normale elenco dei risultati. Che cosa è successo?

Figura 30



Questa variante del nome della cantante è sbagliata, ma è ragionevole pensare che molti utenti nel mondo abbiano commesso la stessa svista. Nella logica del motore di ricerca, un termine sbagliato, quando è molto ricercato, diventa "corretto"!

Questi sono dunque gli effetti del filtraggio collaborativo, la creazione di un sistema in cui i vocabolari sono "entità viventi" originate dai comportamenti collettivi di tutti gli utenti.

La ricerca

Il motore di ricerca

La lingua inganna.

Nel senso di lingua italiana. Quando studiamo il modo migliore di far funzionare il nostro motore di ricerca, ricordiamo che la gente digita indifferentemente parole al singolare e al plurale.

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

Nella maggior parte delle parole, la variante riguarda una lettera: ad esempio "tetto" o "tetti". Ma basta pensare alla coppia "pesca - pesche", per rendersi conto che le lettere che cambiano possono essere due! Se non ne teniamo conto rischiamo di far perdere all'utente tanti risultati.

Esistono anche casi che vanno nella direzione opposta: un termine cercato è contenuto in un altro di diverso significato, e allora qui le risposte rischiano di diventare troppe! Per esempio, "architettura" contiene la stringa "tett" ma in realtà non soddisfa la ricerca sulla parola "tetto".

In questi casi, e soprattutto in riferimento alla lingua italiana, bisogna fornire alcune istruzioni al sistema: avere una tolleranza di due lettere nel confronto delle stringhe di testo, e scartare tutte le altre parole che contengono un numero maggiore di lettere. Con questo stratagemma non raggiungiamo la perfezione, ma abbiamo la possibilità di ampliare la rosa dei risultati senza perdere in qualità. Tutte queste operazioni sono efficaci se svolte in collaborazione con chi controlla la programmazione del codice.

Fieno, papaveri, grano: non tutti i campi sono uguali.

Ce lo hanno spiegato fin dai tempi della scuola: il significato di una parola varia a seconda del contesto in cui la si pronuncia. Anche il peso, la rilevanza sul senso complessivo del discorso più essere diverso da caso a caso.

Ripensiamo alla scheda di contenuto inserita nel sistema di content management. I campi in cui possiamo inserire informazioni sono di tipi diversi: titolo, abstract, testo. Per fare un esempio, il motore di ricerca semplice può segnalare che la parola "tetto" è contenuta in due pagine. In un caso fa parte del titolo ("un nuovo tipo di tetto isolante"), nell'altro è citata en passant in una frase come "il tetto massimo di spesa". Se sono interessato all'edilizia, nel 99% dei casi troverò più interessante il primo documento.

Quindi, più il campo contiene informazioni sintetiche, più il tasso di rilevanza dei termini aumenta. Se la parola cercata coincide con il termine presente in un campo sintetico (abstract, riassunto o sottotitolo), allora significa che siamo in presenza di un contenuto rilevante. Diversamente, occorre fare molta attenzione nel valutare i termini contenuti in campi lunghi e articolati. Si rischierebbe di confondere un termine accessorio, presente quasi per caso, con il centro del discorso.

Per diventare dei veri "information brokers", dobbiamo fornire ai nostri utenti soltanto i contenuti rilevanti.

Le ricerche non devono essere solo parziali, devono dare la possibilità di "vedere tutto"

. Perché le visioni d'insieme sono importanti. Chi ha avuto a che fare con cataloghi aziendali online, sa che immancabilmente arriverà la richiesta: "ma io vorrei vedere tutti i prodotti assieme". Non è ingenuità: la ricorrenza della domanda fa riflettere sulla necessità psicologica di vedere istantaneamente (compatibilmente con il numero complessivo) tutti gli elementi di un catalogo, siano nominativi, aziende, partecipanti, prodotti o altro.

Ecco allora che la proposta di Ipse.com, citata in precedenza parlando di schemi esatti, risponde a questa necessità: io visitatore posso vedere tutti gli elementi in ordine alfabetico, conosco la base documentale. Gli schemi esatti sono molto utili per dare una visione di insieme.

La ricerca vuota: un comportamento tipico.

Chi ha progettato sistemi di ricerca in grandi siti sa qual è la ricerca più frequente: la ricerca vuota. Molti visitatori premono direttamente il tasto "Search" senza digitare nessuna parola. E' l'atteggiamento -legittimo- di qualcuno che vuole sbirciare senza avere nulla in mente. Ma è proprio il risultato di questa ricerca a dover essere curato per primo, quando si pensa a gestire le ricerche nel web. Per convenzione, si risponde ad una ricerca vuota con il valore "tutti i risultati", cioè una lista di titoli completa e non organizzata.

Metadati e presentazione

I metadati e i vocabolari controllati

Quando analizziamo i termini cercati dagli utenti nel motore di ricerca ci accorgiamo che solo una minoranza utilizza parole chiave che definiscono in maniera univoca concetti o argomenti. Questo perché le parole non sono numeri, hanno aloni e somiglianze.

Un buon sistema di ricerca restituisce anche le informazioni che soddisfano i sinonimi della parola cercata. Come è possibile? Attraverso abbinamenti e definizioni stabilite a priori in un Thesaurus. La ricerca diventa così esplorativa; l'utente si muove tra termini diversi dalla parola cercata ma legati ad essa da rapporti di contiguità.

Presentazione dei risultati: il sorting

Il motore di ricerca, costruito con i criteri finora elencati, deve proporre i risultati mettendo in evidenza un ordine di importanza.

Il primo criterio è quello della rilevanza dei campi: prima la pagina in cui la parola chiave è contenuta nel titolo, poi quelle in cui è contenuta nel corpo del testo. Un secondo criterio è quello dei sinonimi: se la pagina contiene due o più termini associati per significato nel thesaurus, allora il grado di pertinenza del risultato è rafforzato.

In ogni caso, un buon sistema di risposta non deve neanche porsi come definitivo. E' importante consentire la possibilità di raffinare la ricerca. Se un utente si accorge che la risposta a quello che stava cercando contiene troppi termini, la possibilità di restringere l'indagine entro i risultati già ottenuti lo aiuterà a chiarirsi meglio le idee.

I visitatori-autori: tecniche di filtraggio collaborativo

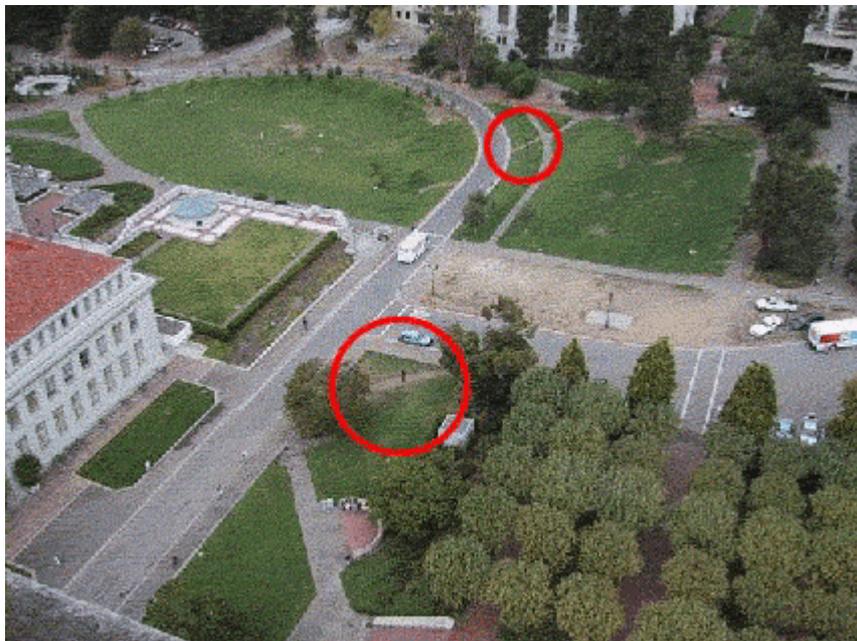
Quando si progetta un sito a tavolino, il bravo webdesigner pensa a tutti i percorsi di navigazione, ai passaggi per ottenere certe informazioni. Un po' quello che succede quando si progetta uno spazio fisico, ad esempio un campus universitario, realizzando un bel plastico ...

Figura 31



Vie di accesso, percorsi guidati, vialetti pedonali... e poi nella realtà cosa succede? Studenti e professori "tagliano gli angoli" e creano scorciatoie!

Figura 32



Il sito web come uno spazio reale, i visitatori creano nuovi passaggi su misura per le loro esigenze. Il filtraggio collaborativo è il processo attraverso il quale i visitatori di un sito contribuiscono a formare le categorie e i raggruppamenti dei contenuti.

Le azioni da tenere sotto controllo sono:

Guida Architettura del sito

Prof. Aldo gorla

sitographics@infinito.it

<http://digilander.libero.it/sitographics>

- la frequenza con cui i visitatori vanno da una pagina A ad una pagina B: una frequenza alta rafforza la validità del link.
- la simmetria: se molti utenti vanno dalla pagina A alla pagina B e molti vanno dalla pagina B alla A, allora il rapporto tra questi due elementi risulta ulteriormente rinforzato.
- la transitività: se l'utente va dalla pagina A alla pagina B e poi dalla pagina B alla pagina C, allora può essere opportuno creare un passaggio da A a C

Metodi di "spionaggio" del visitatore (tracking)

Alcuni grandi siti sono molto interessati a trattenere più informazioni possibili sui loro utenti: quali pagine hanno guardato, che cosa hanno acquistato, da quale nazione provengono, quali prodotti hanno associato. Tutte informazioni che, incrociate fra loro, possono dare grandi risultati dal punto di vista del marketing, portando alla creazione di un vero e proprio Profilo Utente. Non si tratta di una mera profilazione del pubblico di un sito, ma è qualcosa di estremamente più complesso che richiede grandi capacità di analisi e di strutturazione delle basi di dati. Rappresenta la frontiera più avanzata della personalizzazione dei contenuti web e potrebbe aprire nuovi capitoli di studio e discussione. In questa guida ci limiteremo a dire che un filtraggio collaborativo avanzato si basa sulla creazione di un Profilo Utente ottenuto dinamicamente, individuando delle costanti nei percorsi e nei comportamenti dei navigatori

Esempi di filtraggio collaborativo

Un esempio un po' invasivo

L'esempio tipico è Amazon, che nel tempo ha talmente rafforzato il concetto da arrivare a non farci più distinguere tra un oggetto e l'universo dei suoi riferimenti. In altre parole, la pagina di un prodotto inserito nel carrello di Amazon contiene poche informazioni sul prodotto scelto, e in più una miriade di elementi correlati tra loro, scelti attraverso i comportamenti dei visitatori precedenti.

Figura 33



Forse questo tipo di organizzazione della pagina aiuta a vendere, ma è innegabile che generi un po' di confusione.

Un esempio a basso costo

Un esempio di filtraggio collaborativo a costi ridotti, viene dal sito di immagini gratuite stock xchng che tra i criteri di ordinamento dei risultati propone il numero di download che un'immagine ha avuto da parte degli altri visitatori del sito.

In questo modo si determina una lista di importanza che non è predefinita, ma che viene dinamicamente originata dai comportamenti ricorrenti dei visitatori. In ogni caso tutti i siti e portali che adottano il sistema della Link Popularity di fatto forniscono esempi semplificati di filtraggio collaborativo.

Figura 34

