



IA SUMMIT '06
Roma – Aula Convegni CNR
www.iasummit.it

1. **Keynote: User Experience: The next step for IAs?** - P.Boersma (membro della "Board of Advisors" dell'IA Institute nel 2004-2005 ed uno degli organizzatori di EURO IA Summit e dell'IA Summit)
2. **Un percorso a ritroso: le facce nascoste dell'Information Architecture** - L.Caprio & B.Ghiglione (fondatrici della divisione di IA in Inferentia DNM ed attualmente Product Manager in Matrix)
3. **A cosa servono le classificazioni (Classificazioni enumerativa, a faccette e libera)-** C.Gnoli (Dipartimento di Matematica dell'Universita` di Pavia e coordinatore di ISKO Italia)
4. **Modelli di architettura dell'informazione per la Pubblica Amministrazione** - S.Fuchs (Università per Stranieri di Perugia)
5. **Classificare per i cittadini. Il motore di ricerca a faccette della PA piemontese** - M.E. Lai - D.Manfrin (CSI - Piemonte)
6. **Information Architecture per i siti museali: due casi di redesign** - F.Marangoni (Laureata in lettere moderne con Master in Gestione delle Istituzioni Artistiche e Culturali)
7. **Social Distributed Classification e Folksonomies** - E.Quintarelli (Customer Experience Expert in Accenture con Master in Multichannel User Experience)
8. **La classificazione e la struttura relazionale nel Thesaurus: l'esperienza di EARTH** - Plini, Mazzocchi, Di Franco (Istituto sull'Inquinamento Atmosferico CNR)
9. **Abulafia - Sistema di gestione documentale per archivi storico-giudici** - A.Resmini (Dottorando presso il Dipartimento di Discipline Storiche dell'Università di Bologna ed assistente al Politecnico di Milano)
10. **Architettura dell'informazione per il web e mappe concettuali: il caso studio Designing-X** - L.Bollini e G.Palma (Università degli Studi di Milano Bicocca, Dipartimento di Psicologia)
11. **Architettura dell'informazione centrata sull'utente** - S.Bussolon, M.Conci (Scienze della Cognizione e della Formazione, Università degli Studi di Trento)
12. **Blog e Information Architecture** - di T.Siino (Dottorando presso l'Università degli Studi di Palermo)
13. **Information architecture e Interaction design: un modello formativo interdisciplinare** - C.Ferrigno (consulente multimediale Rai), V.Bindi (Sciatto Produzie)
14. **RDF in XHTML: gestire la conoscenza online** - L.Mascaro (Membro IWA/HWG e di diversi gruppi di lavoro internazionali all'interno del consorzio W3C)



Gli interventi

Keynote: User Experience: The next step for IAs? - P. Boersma

Per identificare i next steps dell'IA, Peter Boersma prende spunto dal dibattito internazionale sull'IA e il suo rapporto con le discipline affini, partendo dalla definizione di "Little IA" (il know-how specifico dell'information architecture e dell'information science, relativo principalmente a tecniche di categorizzazione, indicizzazione e information retrieval) e "Big IA" (concepita come l'unione delle competenze specifiche dell'IA e di altri campi affini) data da Peter Morville nel 2004, ed elaborata – anche visivamente – da altri importanti information architects (come JJ Garret).

Peter Boersma pur adottando la stessa differenziazione introdotta da Morville, usando i termini Deep IA e Shallow IA, ritiene che le competenze comuni tra IA e le altre discipline sono identificabili con il concetto e con la disciplina della User Experience (massimo comun denominatore di un vasto ventaglio di discipline) e arriva a rappresentare l'IA con un modello a T, già presentato sul suo blog (<http://www.peterboersma.com/blog>)

L'obiettivo futuro dell'IA non può quindi essere intraviso nella User Experience, ma piuttosto nella necessità di approfondire e lavorare sull'intersezione tra IA e altre discipline (definita con il termine di Guerrilla IA, cioè un insieme di metodi di progettazione semplificati utili per far guadagnare sempre maggior fiducia e credibilità anche alle tecniche più sistematiche dell'IA) definire il campo d'azione e lavorare sulla formazione e sulla pratica (nell'ottica del T-model).

Un percorso a ritroso: le facce nascoste dell'Information Architecture - L.Caprio & B.Ghiglione

Lo spunto di riflessione parte dall'Overload di informazioni che caratterizza l'era digitale, ma si estende anche alle numerose informazioni, personali o meno, che ognuno di noi dissemina su diversi media. Software e standard possono risolvere solo in parte i problemi legati ad una mancanza di organizzazione di informazioni. Si rende necessario un nuovo modo di progettare la strutturazione di spazi informativi.

L'Information Architecture, come disciplina di progettazione centrata sull'utente (UCD) mettendo a fuoco le esigenze degli utenti e gli obiettivi dell'azienda, si concentra sull'organizzazione dei contenuti stessi e sulla progettazione dei sistemi di navigazione e ricerca delle informazioni.

L'Information Architecture si può considerare in diversi modi:

- la metodologia dell'IA: fasi, attività e deliverable
- la grammatica dell'IA: i componenti, le tecniche, gli strumenti
- la pratica dell'IA: le facce nascoste dell'IA nei siti e nei servizi web based.

Essendo, una disciplina ancora poco diffusa, i testi divulgativi e gli articoli dedicati all'Information Architecture spesso si concentrano sui primi due punti di vista, e prendono in considerazione gli aspetti più tecnici della disciplina, focalizzandosi sulle attività e sulle tecniche per svolgere al meglio la progettazione.

In realtà l'Information Architecture non può prescindere dalla definizione del quadro complessivo, cioè degli assi di riferimento che determinano lo spazio in cui un sito web - come fosse un solido poliedrico - prende forma, ovvero contesto, contenuti e utenti.

Si propone quindi di sganciarsi dalla visione tradizionale dell'IA e adottare un punto di vista "rovesciato", che parte dall'utente, dai suoi obiettivi e dalla sua interazione con un sistema, per fare emergere, dietro a semplici esperienze quotidiane, le tracce di un lavoro di progettazione "artigianale", sempre diverso ogni volta.

Prendendo lo spunto da uno scenario complesso - la realizzazione di una intranet- e quindi altamente esemplificativo, vengono messe in evidenza, volta per volta, le facce dell'IA nascoste nel sito, frutto di diverse attività del progettista.



Faccia 1: I contenuti

Lavoro sui contenuti: Strutturare e organizzare i contenuti, creare una tassonomia, progettare la modalità di visualizzazione dei contenuti

Attività correlate: Inventario, categorizzazione, labeling

Altre facce strettamente correlate: Findability, Wayfinding

Faccia 2: i task

Lavoro sui task: semplificare l'esecuzione dei task da parte dell'utente, capire come gli utenti interpretano il funzionamento del sistema (modello mentale), evitare di inventare modelli di interazione nuovi o complessi, rendere soddisfacente l'esperienza dell'utente anche durante la realizzazione del task e non solo al termine

Attività correlate: Personaggi, Scenari d'uso, Task analysis, Task Design, Flussi, Navigazione

Faccia 3: wayfinding

Cos'è il wayfinding: Scegliere e seguire un percorso che porti ad una destinazione definita, in maniera efficiente, Capire dove siamo e dove stiamo andando dalle tracce e dalle piste, Avere il senso della direzione ma anche dello spazio in cui ci troviamo

Attività correlate: Task analysis, Task Design, Flussi, Labeling, Navigazione, Architettura delle pagine

Faccia 4: findability

Attività correlate: Categorizzazione, Labeling, Sistemi di ricerca, Navigazione, Architettura delle pagine

L'individuazione e l'elaborazione di ogni faccia porta a fare luce sulle attività connesse sui relativi deliverables, visti non più in una chiave di lettura metodologica ma come step di un percorso adattivo. Non tutte le attività e tecniche individuate per ogni fase del progetto devono essere rigidamente calate sul progetto, ma solo quelle strettamente connesse e concorrenti alle definizioni delle "facce" del sito o servizio web su cui ci si vuole focalizzare.

A cosa servono le classificazioni (Classificazioni enumerativa, a faccette e libera) - C. Gnoli

Classificazione enumerativa

La maggior parte degli schemi di classificazione sono sostanzialmente enumerativi, ossia si basano sull'elencazione delle classi nelle quali è possibile ripartire i documenti da classificare. Poiché le possibilità di combinare fra loro i concetti sono limitate, quasi tutti i concetti da utilizzare devono essere già previsti dallo schema. L'universo della conoscenza viene suddiviso in un certo numero di classi principali, ognuna di queste in sottoclassi, e così via, sviluppando in questo modo un albero gerarchico, di profondità teoricamente illimitata.

L'esempio più famoso è la Classificazione decimale Dewey (CDD) che utilizza cifre per rappresentare le classi e le loro sottoclassi, analogamente alle cifre dei numeri decimali.

Classificazione a faccette

Un'evoluzione della classificazione enumerativa è la classificazione a faccette. In essa, il contenuto dei documenti è analizzato in una serie di aspetti fra loro complementari, le cosiddette faccette, e quindi espresso per combinazione dei codici corrispondenti a ciascuna faccetta. Un nuovo interesse per la potenza della logica a faccette si riscontra negli ultimi anni nell'architettura dell'informazione. Alcuni siti infatti adottano più o meno esplicitamente un'indicizzazione "a faccette", in cui cioè il contenuto di ciascuna pagina è accessibile alternativamente attraverso l'uno o l'altro degli aspetti che contribuiscono a formarla. Esempi noti di siti organizzati in questo modo



sono Flamenco, InformeDesign e Wine.com. Purtroppo la gran parte dei siti commerciali che si presenta "a faccette" adotta questa tecnica solo in parte: infatti, dopo che i contenuti sono stati scomposti in faccette (analisi), queste non vengono riassemblate secondo un ordine di citazione standard ed espresse conseguentemente da una notazione (sintesi), bensì solo presentate in forma sciolta, elencando i fuochi in ordine alfabetico. Il risultato è una forma ibrida fra classificazione a faccette e semplici parole-chiave.

Classificazione libera

La classificazione libera, a differenza dei due tipi descritti precedentemente, non è ancora stata descritta formalmente in letteratura; tuttavia i suoi principi di fondo sono inconsapevolmente applicati anche in molte basi-dati bibliografiche di letteratura tecnico-scientifica. Anche in questo caso il soggetto di un documento viene scomposto in concetti semplici, i quali però a differenza che nelle faccette "classiche" non sono legati fra loro da rapporti sintattici predefiniti, bensì semplicemente giustapposti. Oltre alle tradizionali ricerche per autore e titolo, gli utenti possono scorrere lo schema di classificazione oppure cercare un termine al suo interno. A ciascuna classe è possibile arrivare anche attraverso diversi tipi di sinonimi e termini associati, grazie a un tesaurus incorporato nello schema. La classificazione libera equivale funzionalmente all'assegnazione di parole-chiave sciolte a ciascun documento: in entrambi i casi, infatti, i componenti semantici sono semplicemente elencati senza che siano specificate relazioni sintattiche. Forme affini alla classificazione libera possono essere considerati anche i sistemi di etichettatura cosiddetti (un po' impropriamente) "a faccette" come quello introdotto da Google Mail, e le folksonomy in cui il contenuto dei documenti è indicizzato mediante termini assegnati liberamente dagli utenti [Quintarelli 2005]. La differenza principale fra questi sistemi e la classificazione libera è che, essendo in quest'ultima i concetti espressi con una notazione invece che con parole, i documenti elencati in ciascun sottomenù risultano automaticamente ordinati in modo sistematico invece che alfabetico, rispecchiando la struttura generale dello schema. Le varie forme di parole-chiave, pseudo-faccette e folksonomy si possono invece considerare come stadi intermedi tra l'indicizzazione alfabetica e la classificazione.

Modelli di architettura dell'informazione per la Pubblica Amministrazione - S.Fuchs

Questo studio tenta di compiere un'analisi sull'efficacia dei criteri di classificazione utilizzati nel portale italiano per il cittadino <www.italia.gov.it>, e in quello britannico <www.direct.gov.uk>. I loro approcci sono totalmente differenti.

Il modello gerarchico di Italia.gov.it: Il governo italiano ha proposto, per organizzare i propri servizi, una struttura gerarchica in cui la via per arrivare all'informazione ambita è una ed una soltanto (eventi della vita). L'adozione della struttura gerarchica genera, secondo la relatrice, numerosi problemi. Uno dei maggiori motivi di difficoltà consiste soprattutto nella rigidità: la via per arrivare all'informazione ricercata è una ed una soltanto, incompatibile con i differenti modelli mentali. Pertanto il modello ad albero o gerarchico si rivela in questi casi inadeguato. Ulteriori problemi sono prodotti dalla sezione principale «Eventi della vita», che è suddivisa in sedici categorie («avere un figlio», «vivere in salute» ecc.) che, però tendono in qualche modo a sovrapporsi tra loro. La presenza de «Le tue guide», una seconda sezione chiave, provoca ancor più confusione. Mediante la differenziazione dei colori (il verde pastello che contrassegna la sezione «Eventi della vita», è sostituito qui da un rosa piuttosto forte), le due sezioni si separano in maniera evidente; ciononostante, non è possibile dire altrettanto a proposito dei contenuti, la cui distinzione non è ugualmente netta. Si potrebbe proporre una struttura alternativa per il portale dei cittadini: la classificazione a faccette. In questo modo i servizi sono esposti da diversi punti di vista e questa organizzazione si rivela pertanto molto più vicina alle capacità e ai criteri della nostra mente. Questo sistema a faccette ammette differenti interpretazioni dello stesso servizio e, di conseguenza, numerosi modelli mentali.

Il modello associativo di Direct.gov.uk: Due anni fa è stato inaugurato il nuovo portale della Pubblica Amministrazione britannica, battezzato col nome di Directgov <www.direct.gov.uk>.



Abbandonando il modello gerarchico e usando un sistema poligerarchico, (un sistema non ad albero che offre all'utente numerosi percorsi per arrivare ai servizi) i servizi sono raggiungibili attraverso molteplici percorsi. Questa qualità del sito si manifesta esplicitamente nelle due sezioni primarie: «Straight to...» (che classifica il contenuto mediante i vari argomenti) e «People» (che distingue il contenuto per tipi di utente); a queste due vengono associati ulteriori punti di partenza, come «Popular now», «Top choices», «Do it online» and «Guide to govt».

Molti utenti sono abituati a dare una rapida occhiata alla pagina su cui si trovano, per poi cliccare sul link che, meglio degli altri, cattura la loro attenzione; per questo motivo i creatori hanno impostato le procedure di navigazione proprio sulle parole chiave. Tutte le categorie delle due sezioni «Straight to...» e «People» sono elencate sull'homepage e anche una parte delle loro sottocategorie. In questo modo le categorie diffondono un ricco e ben distinguibile "information scent"; grazie a questo, poi, l'utente riesce ad anticipare mentalmente le pagine successive e ad immaginare possibili percorsi.

Classificare per i cittadini. Il motore di ricerca a faccette della PA piemontese - M.E. Lai - D.Manfrin (CSI - Piemonte)

Il SistemaPiemonte è il portale territoriale della PA piemontese, www.sistemapiemonte.it che costituisce il punto di accesso unificato a servizi/informazioni, si rivolge a cittadini, professionisti, imprese e offre servizi interattivi (es. ricerca e visualizzazione dati, download modulistica, inoltro pratiche, esecuzione transazioni).

Obiettivi del progetto sono soddisfare bisogni e rispettare aspettative e aumentare la ritrovabilità. Per soddisfare questi obiettivi, è stato scelto uno strumento specialistico, efficiente e usabile: un motore di ricerca che si basa su una tassonomia a faccette e sull'utilizzo di classificatori semantici automatici capaci di ricondurre i contenuti on line alla tassonomia individuata.

La classificazione a faccette è stata adottata innanzitutto per superare i limiti delle tassonomie tradizionalmente utilizzate sui siti degli Enti pubblici e per proporre, attraverso un approccio multidimensionale, molteplici chiavi di accesso alle informazioni nel rispetto delle esigenze e delle aspettative degli utenti. L'introduzione del classificatore automatico è resa necessaria dall'ampiezza dei domini da analizzare e dall'impossibilità di tenere sotto controllo gli aggiornamenti in tempo reale da parte della redazione. Il motore prevede differenti modalità di ricerca per consentire agli utenti di adottare le strategie più consone alle abitudini e al grado di conoscenza degli argomenti. Accanto alla tradizionale ricerca per parole chiave, grazie ai risultati delle classificazione multidimensionale a faccette fornita dal classificatore automatico, è possibile effettuare ricerche direttamente sulla tassonomia, seguendo un percorso di navigazione a discesa sull'albero delle faccette. La ricerca attraverso l'esplorazione della tassonomia ha lo scopo di agevolare quegli utenti che si accostano al contenuto informativo di un sito senza avere un'idea chiara di ciò che vogliono o possono trovarvi, e che quindi - in una prospettiva esplorativa - gradiscono poter "dare uno sguardo" a ciò che è disponibile.

Inoltre il motore si avvale del vocabolario controllato, che ha lo scopo di allargare il dominio di ricerca e di introdurre un livello interpretativo.

L'attività di ricerca sulla terminologia ha permesso di abbozzare la struttura del vocabolario controllato, considerato come punto di incontro tra lessico dell'autore e lessico degli utenti, e come mezzo per superare le difficoltà legate all'uso del linguaggio specialistico (a ciascun termine è associata una serie di sinonimi, locuzioni affini e, se possibile, rimandi logici ad esempio: ticket → compartecipazione alla spesa sanitaria).

I termini inseriti sono stati selezionati in base ad analisi di report statistici dei motori di ricerca interni ai siti del Comune di Torino e di Sistema Piemonte che evidenziano le espressioni più ricorrenti; consultazione di risorse linguistiche ed un esame di un campione rappresentativo di documenti disponibili in rete.



La progettazione del portale territoriale ha seguito per tutta la durata del progetto i principi dell'UCD, prevedendo momenti successivi di verifica e aggiornamento:

- raccolta preliminare di documenti sulle best practices
 - o analisi dell'interazione su servizi simili: es. Flamenco, Sealink
- interviste strutturate con gli utenti (maggio 2005)
 - o team composto da un esperto di fattori umani (M. Boscarol), progettisti
 - o utenti che interagiscono con la PA e ricercano informazioni o servizi online
 - o domande sull'uso abituale del web, sulle strategie di ricerca, sull'interazione con la PA. Brevi test di verifica.
- test su prototipi interattivi e cartacei (rapid prototyping)
- analisi dei risultati e realizzazione del prototipo
- test di usabilità, card sorting

Al momento, è stato sviluppato un prototipo del motore, sia lato front-end che back-end.

Le funzionalità associate alla componente di front end sono:

- ricerca testuale libera
 - o analisi morfologica della query utente → dizionario elettronico
 - o analisi sintattica per riconoscere espressioni → es. "settembre musica"
 - o uso del "forse cercavi..."
 - o creazione dell'abstract
- navigazione sulla tassonomia
 - o si visualizzano solo le classi contenenti elementi. Non ci sono classi vuote.
 - o selezione oppure esclusione sequenziale delle classi.
- ricerca mista
 - o uso combinato di *search* e *browsing* per raffinare il risultato
- valore aggiunto: il vocabolario controllato stabilisce corrispondenze fra le parole inserite dall'utente e quelle presenti nei documenti (es: digitando "ticket" il sistema reperisce tutti i documenti relativi alla "compartecipazione alla spesa sanitaria" (terminologia burocratica)

Information Architecture per i siti museali: due casi di redesign - F.Marangoni

L'intervento mostrare come sono stati applicati ai due progetti alcuni principi e metodologie di design dell'architettura informativa, evidenziando in particolare le attività e le soluzioni considerate particolarmente strategiche per gli obiettivi richiesti da ciascun sito.

Per entrambi i siti (<http://www.museostorico.tn.it> e <http://www.buonconsiglio.it>) la situazione attuale è quella di un'impostazione superata, con forti problemi di usabilità, navigabilità e reperibilità delle informazioni. Per entrambi si pone l'esigenza di ampliare e ridefinire i contenuti, di dotarsi di una architettura informativa più flessibile e user oriented, basata su un modello relazionale e che possa sostenere ulteriori sviluppi e ampliamenti, nella prospettiva della costruzione di un "portale di servizi" museale, capace di porsi come canale dell'offerta culturale complementare a quelli tradizionali ed è necessario per i due siti un nuovo design visuale, che sia progettato secondo gli standard di usabilità dei siti web.

Nella soluzione individuata si è ritenuti indispensabile porre al centro della progettazione il design dell'architettura informativa, individuando alcuni aspetti e attività considerati strategici.

Una volta censito l'inventario dei contenuti, per entrambi i progetti è emersa l'esigenza di classificare i contenuti secondo un approccio multidimensionale, che permettesse:

1. lo sviluppo di un sito basato prevalentemente su una struttura relazionale (modello database)
2. lo sviluppo di un'interfaccia utente che potesse garantire un accesso multidimensionale e circolare all'informazione, attraverso un'attenta progettazione del sistema di navigazione (soprattutto della navigazione contestuale) e degli strumenti di ricerca e selezione



Ai primi livelli della mappa del sito, una struttura gerarchica definisce il menu di navigazione globale e locale:

- nel livello più alto (navigazione globale) la classificazione è basata su ambiti o domini dell'offerta culturale, nei quali gli utenti riconoscono l'organizzazione dei tradizionali servizi e attività di un museo.
- nei livelli immediatamente inferiori (secondo e terzo), raggiungibili attraverso il menu di navigazione locale, i servizi e le attività sono organizzati in base a criteri non omogenei per tutte le aree ma sempre secondo uno schema organizzativo esatto: così, ad esempio, le mostre sono suddivise in temporanee e permanenti (secondo uno degli standard più comuni a tutti i musei), mentre i servizi formativi sono classificati in base alla tipologia dei destinatari.

Una delle caratteristiche che il modello ancora non supporta efficacemente e che si vorrebbe implementare riguarda la possibilità di classificare alcuni contenuti in base agli argomenti. I contenuti in questo caso sarebbero selezionati tra quelli che si suppone possano essere oggetto di ricerca da parte di utenti che sanno cosa cercare e che eventualmente conoscono il dominio: pubblicazioni e indici delle riviste, seminari e incontri pubblici, progetti di ricerca, attività espositive, attività educative.

La classificazione per argomenti è uno dei più tipici e comuni casi di classificazione ambigua, ma anche uno dei più utili. Questo tipo di classificazione deve essere definito in base alle aspettative degli utenti (e di diversi tipi di utenti), favorendo una modalità di ricerca tipicamente esplorativa. Nel nostro caso le problematiche relative alla classificazione per argomenti sono le stesse già evidenziate dai due autori e sono dettate prevalentemente dall'eterogeneità degli utenti e dall'ambiguità nell'uso dei termini. Volendo garantire la reperibilità dell'informazione in uno schema organizzativo di questo tipo, la soluzione ottimale è rappresentata dal thesaurus classico, caratterizzato da relazioni di equivalenza, gerarchiche e associative.

Punti di Forza

I metodi di progettazione adottati hanno permesso di sviluppare un modello di sito che potenzialmente garantisce l'accesso multidimensionale all'informazione e quindi contribuisce enormemente al miglioramento della navigabilità e dell'usabilità del sito, attraverso l'integrazione tra classificazione gerarchica dei contenuti (navigazione globale e locale) e modellazione dati di tipo relazionale (navigazione contestuale interna alla pagina e strutturata).

La modellazione dei dati (ERD) permette la flessibilità e l'estensibilità del modello. Con un'appropriata attività di analisi e classificazione è possibile scegliere i criteri descrittivi (per target d'utenza, per tematiche, per tipologia di attività, contenuto o servizio) sui quali basare di volta in volta le strutture organizzative da applicare alla navigazione, all'architettura della pagina e alle zone di ricerca, eventualmente estendendo il diagramma originario.

Criticità

Le criticità sono riconducibili prevalentemente al lavoro vero e proprio dell'architetto dell'informazione e cioè all'analisi e alla classificazione dei contenuti in base agli obiettivi che gli utenti devono poter raggiungere navigando nel sito.

Più in particolare si rende necessaria un'analisi con gli utenti per definire schemi di classificazione ambigui che possano supportare la ricerca (e l'eventuale navigazione) per argomenti, considerando la complessità già descritta per il caso del Museo Storico e che potrebbe facilmente ripresentarsi in altri casi.

Più prematura appare invece la possibilità di stabilire classificazioni ambigue in base al target di utenza, sia per la difficoltà di svolgere indagini approfondite presso gli utenti, sia per i problemi intrinseci di questo criterio di organizzazione.



Social Distributed Classification e Folksonomies - E.Quintarelli

Negli ultimi tempi stiamo assistendo alla creazione contenuti web senza precedenti (web log, wiki, social tools), che non sono generati da utenti esperti e professionisti dell'informazione ma da persone al tempo stesso utenti e produttori di informazioni.

Questa incredibile mole di informazioni devono in qualche modo essere aggregate e categorizzate. Gli schemi tradizionali di classificazione funzionano bene in presenza di alcune condizioni: il contenuto deve essere relativamente ristretto, deve avere categorie stabili e contenere elementi omogenei. La classificazione, inoltre, deve essere svolta da professionisti ed è rivolta ad utenti esperti.

La realtà che si sta ora delineando presenta caratteristiche del tutto diverse: fonti eterogenee producono una mole enorme di informazioni, in continua evoluzione e non chiaramente definite, rivolte ad un target altrettanto eterogeneo: questa situazione richiede nuove strategie di classificazione.

Le folksonomies possono essere considerate come un approccio innovativo. Il termine folksonomy (Folk + Taxonomy) indica una pratica di categorizzazione collaborativa, realizzata dagli stessi utenti tramite l'utilizzo e attribuzione di parole chiave scelte liberamente.

Questo permette agli utenti di condividere le proprie parole chiave e di far emergere in maniera naturale associazioni tra contenuti e parole chiave.

Alcuni tra i primi siti ad introdurre questa modalità di categorizzazione sono stati **Del.icio.us** (<http://del.icio.us>), **Flickr** (<http://www.flickr.com>), **43things** (<http://www.43things.com>), **Furl** (<http://www.furl.net>), **Technorati** (<http://www.technorati.com>), che offrivano la possibilità di utilizzare parole chiave (o tag) come feature.

L'utilizzo delle folksonomies non è privo di difetti: mancanza di precisione, classificazione non gerarchica, bassa possibilità di reperimento. Tuttavia, i pro sono notevoli: buon compromesso tra semplicità e precisione, vicinanza con il linguaggio degli utenti e le sue esigenze, stimolano la serendipity, permettono ad ogni utente di far sentire la propria voce e, in ultimo, soprattutto per contenuti particolari che rischiano di non essere altrimenti reperibili (es. fotografie, link) sono meglio di nulla. Per tratteggiare uno scenario futuro, di possibile evoluzione delle folksonomies, Quiinaterlli ipotizza tre strade per migliorare il tag: tag a faccette (es. <http://www.mefedia.com/tags/>), navigazione avanzata nei tag (<http://johnvey.com/features/deliciousdirector/>) e tag gerarchici (<http://resource.smartdesktop.org/rescon/>).

La classificazione e la struttura relazionale nel Thesaurus: l'esperienza di EARTH - Plini, Mazzocchi, Di Franco

Il CNR ha avviato un progetto relativo allo sviluppo di un thesaurus generale per l'ambiente (EARTH).

Il thesaurus dovrà essere uno strumento ben strutturato e rifinito, rappresentare una mappa semantica e terminologica aggiornata del dominio ambientale, permettere vari livelli di comprensione e applicazione per utenti con differenti caratteristiche ed esperienza e assicurare l'esportazione del thesaurus in differenti applicazioni tecnologiche.

Lo schema di classificazione di EARTH è basato su un sistema di categorie.

Seguendo una prospettiva dal basso verso l'alto, i termini possono essere analizzati secondo una scala progressiva gerarchica. Si raggiunge così il massimo livello di genericità.

La struttura verticale del Thesaurus è fondata sulle categorie e comprende le relazioni gerarchiche. Essa funziona come uno strumento che si propone di guidarci verso i "significati primari" delle cose, almeno per come sono intesi secondo la logica interna al sistema e può fungere da riferimento semantico stabile e, almeno parzialmente, indipendente dal contesto.



Il modello progettato prevede anche che tale struttura di base possa essere complementata da un sistema di temi, che incrociandosi con la struttura verticale andrebbe a costituire un sistema a matrice. Il sistema di temi, per come è stato concepito, dovrebbe essere sviluppato ogni volta in rapporto alle esigenze specifiche del contesto applicativo, che definisce appunto ciò che è tema, cioè oggetto di interesse. Nella lettura per temi, vengono riaggregati i termini associati ad un determinato settore (tematico), termini che la struttura a faccette, collocandoli sotto il concetto più generale di riferimento, tende a sparpagliare. La possibilità, inoltre, di applicare diverse letture tematiche crea le condizioni per analizzare i concetti sulla base di diverse prospettive, che possono evidenziarne aspetti particolari e contingenti.

Il modello consente la possibilità di sviluppare una ulteriore organizzazione della terminologia. La struttura verticale può essere completata da un "micro-mondo" di termini connessi tematicamente (temi).

Mentre la struttura ad albero tende a spargere i termini nelle loro categorie di riferimento, i temi accorpano i termini secondo la loro prospettiva espressa dai temi stessi.

I temi in EARTH forniscono una prospettiva aggiuntiva per l'interpretazione del termine e agiscono come strumenti per rappresentare altri tratti semantici.

Tema SALUTE → *benzene* come sostanza tossica.

Tema INQUINAMENTO → *benzene* come inquinante.

Tema SICUREZZA → *benzene* come sostanza pericolosa.

Questo modello deve anche permettere la rappresentazione di un significato secondo accezioni secondarie.

I thesauri tradizionali forniscono una serie limitata di relazioni fra i termini, distinguendo solo tra relazioni gerarchiche, relazioni associative e relazioni di equivalenza.

Inoltre le relazioni nel thesaurus sono spesso applicate in modo incongruo. Questo causa ambiguità nell'interpretazione e può dare luogo a strutture semantiche imprevedibili.

Nel progetto di sviluppo del Thesaurus è prevista una differenziazione e migliore espressione semantica delle relazioni, e in particolare verrà rinforzata, la struttura trasversale delle RT (related terms). L'arricchimento e l'esplicitazione logica delle relazioni, ne rafforzano la potenzialità di utilizzo per la navigazione su base concettuale e ne aumenteranno le possibilità di utilizzo e riuso in diversi ambiti applicativi. Da un altro punto di vista, l'incremento e l'articolazione delle relazioni associative potenzieranno la capacità del sistema di rappresentare l'area concettuale nella rete di interconnessioni che la caratterizzano e, in termini applicativi, di migliorare la risposta alle esigenze poste dalla ramificazione dell'informazione su Internet.

Un altro obiettivo da raggiungere è assicurare una alta modularità del sistema.

Non tutti gli utenti sono interessati nelle sottili distinzioni delle relazioni del thesaurus.

Sarà possibile navigare nella struttura del thesaurus per differenti livelli, a partire dalla versione tradizionale della struttura relazionale del thesaurus.

Abulafia - Sistema di gestione documentale per archivi storico-giudici -

A.Resmini

Partendo dal paradosso informativo (definito come infoglut), per cui effettuando una ricerca quanti più risultati trovo tanto più difficile è trovare informazioni rilevanti, filtrare il rumore, poterne valutare autorevolezza e validità temporale, emerge chiaramente la necessità di migliorare la cercabilità nelle applicazioni che gestiscono dati sul web.

Tra i rimedi esistenti messi in campo possiamo enumerare la ricerca full text, l'aggiunta di meta-data esterni, l'implementazione delle faccette, di RDF e altri standard.

Un altro rimedio possibile consiste nell'utilizzo delle topic maps.

Le topic maps sono definite da uno standard ISO (ISO/IEC 13250) e sono un sistema per l'organizzazione e la correlazione di dati indipendente dai dati stessi.



Le topic maps realizzano un indice che risiede al di fuori delle informazioni che descrive: si lavora con il significato delle informazioni e non più con le informazioni stesse. Il risultato è una struttura informativa non vincolata dai tradizionali sistemi gerarchici, che contiene gerarchie multiple sovrapposte semanticamente ricche e che consente all'utente di non dover seguire più la sola logica di accesso pensata dal progettista: i percorsi sono multipli e ridondanti. Le topic maps non obbligano ad un modello ontologico predefinito (sono in effetti meta-modelli), ma consentono di definire il proprio sistema organizzativo per l'architettura dell'informazione modellandolo come una classificazione a faccette, un vocabolario controllato, un tesauro, un'ontologia o una combinazione di questi.

Una topic map è costituita principalmente da topics. Questi rappresentano le cose descritte dalla topic map. Adattando un esempio di L. M. Garshol, in una topic map che tratti di IA possibili topics sarebbero deliverables, usability heuristics, Nielsen, o Morville.

I topics hanno relazioni, che la topic map modella come associazioni (associations) e che sono typed, ovvero descritte. Seguendo il nostro esempio, la association tra Nielsen ed usability heuristics è una relazione di authorship. Le associazioni hanno una caratteristica interessante: i topics all'interno di una relazione ricevono un ruolo (role). Ad esempio, Nielsen è l'autore (author), mentre le heuristics sono il lavoro (work).

Il significato di questo semplice passaggio è che all'interno di una topic map le due frasi 'Nielsen ha scritto le usability heuristics' e 'Le heuristics sono state scritte da Nielsen' sono esattamente la stessa cosa, l'associazione tra i topics è bidirezionale.

L'ultimo elemento di una topic map sono le occorrenze (occurrences), ovvero risorse informative rilevanti in relazione ad un topic. Nel caso di Nielsen, una foto, la sua homepage, un saggio. Poiché anche le occorrenze possono essere descritte (typed), le informazioni riportabili ad un utente che richiede un topic non sono solo quantitative (un certo numero di link), ma anche qualitative (cosa sono quei link).

Con un vocabolario ristretto, elegante e semplice, le topic maps possono generare un'elevata complessità semantica.

I CMS hanno risolto diversi problemi legati alla lavorazione di contenuti in ambito multi-utente, lasciando però aperte diverse questioni. Tra queste, non ultima, l'organizzazione delle informazioni e la conseguente reperibilità da parte di una platea eterogenea di utenti non noti a priori.

I CMS hanno sviluppato propri modelli tassonomici ad hoc, diversi tra loro, risolvendo il problema specifico ma irrigidendo su questi le proprie funzionalità ed imponendo limitazioni alla disponibilità delle informazioni. Le topic maps forniscono il tassello mancante: un modello descrittivo infinitamente flessibile, ma standard.

All'interno di un CMS le topic maps lavorano a più livelli: producono la navigazione, consentono la cercabilità, incrementano la findability, creano parte del contenuto attraverso le occorrenze dei propri topics.

Abulafia è un progetto di gestione documentale per archivi storico-giuridici, con uno schema di metadati basato sul modello elaborato dal Dublin Core (DCES).

Il progetto e l'implementazione di Abulafia consisterà in due step principali: realizzazione della topic map e della base dati (raccolta e disamina delle fonti, definizione di vocabolario controllato e di un'ontologia, implementazione della base dati TMRQL, popolamento dell'ontologia, ovvero costruzione della topic map vera e propria, implementazione della base dati SQL e popolamento della base dati, ovvero costruzione dell'archivio informativo) e realizzazione dell'applicazione (implementazione dell'applicazione che utilizza la topic map per rendere disponibile i dati attraverso views e stored procedures TMRQL e realizzazione di back-end e front-end per la gestione e la consultazione dei dati)

Nel caso un progetto come Abulafia, che nasce per risolvere un problema specifico relativo ad una base documentale specifica e a scopi sperimentali, molto del lavoro di costruzione semantico della prima fase, inclusa l'analisi delle fonti per la creazione di *topics*, *types*, *associations*, è in un certo qual modo facilitato dall'omogeneità dei dati di partenza, mentre, anche considerando il contesto sperimentale, l'implementazione di una topic map TMRQL e la necessaria costante verifica ed affinatura di questa implementazione presentano sicuramente più problemi progettuali.



Architettura dell'informazione per il web e mappe concettuali: il caso studio Designing-X - L.Bollini e G.Palma

La navigazione nello spazio ipertestuale del web prevede che l'utente abbia degli strumenti concettuali e visivi che gli permettano di orientarsi nella ricerca di informazioni e di crearsi un modello mentale della struttura di un sito per capire e apprendere il funzionamento.

Tra gli strumenti che nell'architettura dell'informazione si usano per rappresentare relazioni e percorsi, le mappe sono un modo per visualizzare spazialmente le informazioni e per veicolare maggiori dettagli in maniera contemporanea anziché sequenziale.

Le mappe cognitive sono uno strumento per rappresentare il sapere e per aiutare ad apprendere e a memorizzare ciò che si impara. Le mappe sono articolate secondo nodi/unità semantico-concettuali interrelati da link/preposizione e possono essere utilizzate per rappresentare la conoscenza in una dinamica non lineare del processo di apprendimento.

La struttura reticolare delle mappe è molto simile a quella delle strutture informative ipertestuali basati su unità informative o nodi costituenti unità comunicative e di significato autonome e le loro correlazioni.

L'uso delle *mappe cognitive*, come sistema di rappresentazione nelle interfacce web permette di interpretare e soddisfare le logiche associative di un utente in maniera spesso più efficace rispetto alla tradizionale struttura ad indice utilizzata nei siti, grazie ad una lettura immediata di molti elementi informativi e delle loro relazioni. Le mappe sono quindi ottimi strumenti per interpretare, rielaborare e trasmettere conoscenze, informazioni e dati, visualizzando l'oggetto della comunicazione, i concetti principali, i legami che essi stabiliscono e, di conseguenza, il percorso del ragionamento.

Il sito Designing-x rappresenta un progetto sperimentale per la realizzazione di un sito di tipo cognitivo-esperienziale.

Il sito è un *contenitore* di cinque realtà diverse legate al mondo del design. Il sito, infatti, riunisce in sé contenuti, servizi, basi di dati, pubblicazioni, ricerche, in un discorso omogeneo, strutturato, complesso. Lo scopo di *designing-x* è sviluppare in un progetto coerente tutti i servizi offerti da questi attori in una modalità *di relazione*. I contenuti trattati verranno presentati come parte di un'unica offerta evidenziando non tanto la suddivisione tra le sezioni e il relativo sviluppo in contenuti, ma la relazione esistente tra i contenuti appartenenti alle diverse sezioni.

Questa scelta si giustifica con la necessità di rendere immediatamente percepibile la complessità della strutturazione dei contenuti nelle sezioni e tra le sezioni, garantendo allo stesso tempo una profonda semplicità di navigazione e di fruizione.

Architettura dell'informazione centrata sull'utente - S.Bussolon, M.Conci

L'Università di Trento ha deciso di riprogettare il proprio sito web in base ad obiettivi organizzativi (es. decentramento e autonomia nella gestione dei contenuti), di usabilità e accessibilità (es. adeguamento alla normativa in vigore sull'accessibilità).

L'intervento di Bussolon e Conci nella ristrutturazione di UNITN è focalizzato alla ridefinizione dell'architettura dell'informazione.

In una prospettiva centrata sull'utente la struttura informativa dovrebbe essere basata sui modelli mentali degli utenti, ed è sulla categorizzazione che si è concentrato l'esperimento.

La metodologia utilizzata si basa su quella sviluppata da Sinha and Boutelle, per creare un'architettura dell'informazione user-centered e future-oriented. Prevede tre fasi: esplorazione delle informazioni del dominio, comprensione degli obiettivi del business ed il contesto (analisi degli stakeholder e free listing), elicitazione delle categorie implicite degli utenti attraverso il card sorting aperto, elicitazione della categorizzazione delle voci nelle categorie ottenute attraverso il card sorting chiuso



Attraverso il free listing è possibile generare una lista esaustiva delle voci di un dominio semantico. La lista è, ragionevolmente, rappresentativa sia delle aspettative degli utenti (ovvero di ciò che gli utenti si aspettano di trovare nel sito) che della terminologia da essi utilizzata.

Il free listing viene utilizzato prima di procedere con il card sorting, per determinare il dominio semantico, ovvero le voci da utilizzare nel card sorting.

Il card sorting è la tecnica di elicitazione della conoscenza più usata e citata nell'area dell'Interazione Uomo computer per far emergere i modelli mentali degli utenti sulle aspettative di categorizzazione della struttura di un sito web.

Il card sorting può essere somministrato in due modalità: card sorting aperto e card sorting chiuso. Nel card sorting chiuso all'utente viene chiesto di categorizzare gli items in categorie stabilite dallo sperimentatore. Il card sorting aperto è meno strutturato in quanto è l'utente che decide il nome delle categorie, e dunque permette di far emergere i criteri di categorizzazione impliciti degli utenti. I risultati del c.s. aperto sono particolarmente interessanti poiché possono darci informazioni su delle tipologie di categorizzazione che non immaginavamo, anche se la maggiore arbitrarietà concessa all'utente rende meno coerenti i risultati raccolti.

Per l'applicazione dei test di free listing, card sorting aperto e chiuso è stato utilizzato Netsorting, una web application che permette di utilizzare – on line – differenti metodi di elicitazione della conoscenza: categorizzazione via card sorting (aperto e chiuso), valutazione di importanza: ai partecipanti viene richiesto di valutare la prototipicità di una lista di items in una categoria, free listing (generazione di items): si chiede ai partecipanti di produrre degli items data una categoria. Nella prima fase dell'esperimento è stato usato il free listing per ottenere la lista degli items da somministrare nel card sorting. Nella seconda fase dell'esperimento abbiamo applicato il metodo del card sorting aperto alle 120 voci ottenute dal free listing. In questa fase abbiamo ottenuto una classificazione preliminare delle voci e una lista di categorie che stiamo utilizzando nel card sorting chiuso. Nella terza ed ultima fase dell'esperimento stiamo utilizzando la tecnica del card sorting chiuso. La procedura sperimentale utilizzata è la stessa della fase precedente, con la differenza che agli utenti viene chiesto di categorizzare l'elenco degli items in categorie già etichettate da noi. La scelta delle categorie-contenitori è stata fatta in base sia ai risultati del card sorting aperto che alle esigenze espresse dal team di UNITN.

Tra i problemi riscontrati durante l'esperimento, è da segnalare un alto tasso di abbandoni e difficoltà nella classificazione di alcuni item.

Blog e Information Architecture - di T.Siino

Per blog si intende «un sito web che, riprendendo alcuni aspetti del diario cartaceo e beneficiando della multimedialità, contiene frasi, articoli, racconti, poesie, foto, video e collegamenti organizzati in ordine cronologico (dal più recente al meno recente) e aggiornati di frequente. Non tutti gli elementi elencati devono essere necessariamente presenti. Il blog può essere a ragione definito un «diario virtuale», ma tale definizione rischia di creare un equivoco e di far sì che i blog vengano ricollegati ai diari di vita (o adolescenziali). È maggiormente corretto, invece, parlare di diario nel senso di «diario intellettuale». Blog è la forma contratta di weblog, che letteralmente significa «registro cronologico di ciò che avviene, riportato sul web».

Secondo il relatore, l'aumento esponenziale del numero dei blog attivi deve essere analizzato anche dal punto di vista dei benefici apportati alla user experience nella navigazione e nella ricerca delle informazioni. I blog presentano una disposizione delle informazioni nella pagina abbastanza standardizzata e ormai riconoscibile ai fruitori di Internet. Alcuni spill-over della grande diffusione del modello blog riguardano la diffusione delle folksonomy e la classificazione delle informazioni con meta-informazioni.

I blog, quindi, possono essere navigati anche in base a percorsi basati su una singola parola chiave. L'interfaccia tipica rispondente a questo criterio è un modulo denominato tag cloud o weighted list. Si tratta di una lista dei tag utilizzati su un sito o un blog che enfatizza con la grandezza e l'eventuale applicazione di un stile grassetto ai tag o alle parole più utilizzate.



Information architecture e Interaction design: un modello formativo interdisciplinare - C.Ferrigno V.Bindi

L'esperienza inizia nel 2003-2004, quando i relatori sono stati coinvolti nella prima edizione, sperimentale, di un master in progettazione multimediale.

La forza del modello di interazione tra architettura dell'informazione e design dell'interazione è stata anche quella di avere a disposizione un progetto "vero", cioè la realizzazione di un progetto di sito per la futura Città della Scienza di Roma, commissionato dalla Direzione Musei Scientifici della Sovrintendenza Comunale ai Beni Culturali di Roma.

Il primo passo è stato quello della raccolta dei dati, che ha coinvolto dal vivo anche il cliente stesso, fornendo materiale sufficiente per riflessioni approfondite sul target.

Il lavoro dei ragazzi è stato quello di adottare la tecnica delle Personas di Alan Cooper, per costruire personaggi e scenari.

Negli anni successivi il corso è stato "sistematizzato", affinando il lavoro interdisciplinare delle due cattedre. Il lavoro in laboratorio continua ad essere impostato su Personaggi e Scenari e a vertere su progetti concreti, ma sono state rimodellate in parte anche le cattedre correlate, dando più spazio all'accessibilità e all'usabilità, ridefinendo il ruolo della multimedialità, scegliendo la strada dell'educazione al linguaggio delle immagini, in particolare da utilizzare sul web; aggiornando l'insegnamento dell'html. Si è anche scelto di assegnare ad ogni componente dei gruppi di lavoro un ruolo esplicito (capo progetto, responsabile dell'architettura informativa, responsabile dei contenuti, responsabile della grafica, responsabile dello sviluppo): che rappresenta uno step importante nel percorso formativo, perché pur venendo "allevati" nello spirito di un team interdisciplinare ognuno di loro deve anche imparare a rispondere in prima persona delle proprie scelte e decisioni.

Pur essendo consapevoli che si tratta di una piccola esperienza, molto artigianale, c'è la piena convinzione che la progettazione centrata sull'utente sia uno strumento validissimo nel campo professionale, ed eccezionalmente creativo dal punto di vista didattico.

RDF in XHTML: gestire la conoscenza online - L.Mascaro

Il web è stato pensato per contenere un infinito numero di informazioni e dati possibilmente tipizzati e relazionati tra loro secondo i principi dell'ipertesto senza però avere dei reali meccanismi di gestione della conoscenza o di inferenza sugli stessi dati.

Il web semantico nasce per coprire questa lacuna del web normale fornendo meccanismi quali le triple o le ontologie per relazionare più profondamente i dati e le informazioni per generare della reale conoscenza.

Purtroppo la difficile applicabilità su larga scala del web semantico al momento lo racchiude unicamente in sistemi chiusi facilmente replicabili con altre tecnologie.

Ma come gestire la conoscenza su larga scala? Alcuni propendono per una fantomatica semantica di XHTML ma con l'avvento di XHTML, specialmente alla versione 2.0, si prospettano nuove modalità per dare un vero significato ai nostri contenuti.

Cos'è un dato? Cos'è un'informazione? Cos'è la conoscenza?

Ad oggi chi si occupa di gestione della conoscenza nell'informatica identifica in genere le seguenti tre definizioni (semplificate):

Dato: Un dato è ciò che è immediatamente presente alla conoscenza e su cui un'elaborazione mentale può portare ad un'informazione.

Per esempio: Data: 24.02.2006 – Città: Roma – Temperatura: 17 - scala: centigradi

Informazione: Un'elaborazione di dati atta a comprensione umana degli stessi.

Per esempio: "Il 24 febbraio 2006 nella città di Roma la temperatura era di 17 gradi centigradi"

Conoscenza (wikipedia): La conoscenza è la consapevolezza e la comprensione di fatti, verità o informazioni ottenuti attraverso l'esperienza o l'apprendimento (a posteriori), ovvero tramite l'introspezione (a priori).



Oggi gli strumenti teoricamente più consoni a rappresentare la gestione della conoscenza sono le basi di dati; piccoli o grandi contenitori di dati strutturati, classificati e relazionati tra loro su cui è possibile effettuare delle elaborazioni (inferenze logiche).

Il problema principale di cui soffrono le basi di dati è che sono principalmente sistemi chiusi su schemi logici ben definiti ed in genere non condivisi. Inoltre non contengono informazioni ma dati puri che vanno rappresentati per ottenere delle informazioni e quindi della conoscenza.

Per rappresentare questa conoscenza ci serviamo di sistemi software che si occupano di elaborare questi dati, estrapolandone dei nuovi e inferendo su di essi, con i quali infine rappresentiamo, in genere visualmente, delle informazioni.

Questi sistemi software sono tipicamente sistemi chiusi che non condividono le loro capacità con il resto del mondo ma, in alcuni casi, ne rappresentano le informazioni tramite le tecnologie web.

Sulla base della premessa precedente si può comprendere come il web in realtà non è un sistema software che contiene dati ma è un sistema che rappresenta informazioni.

Per questo motivo abbiamo grandi vantaggi a livello di diffusione della conoscenza ma i motori di ricerca o gli altri strumenti software, che utilizzano il web come fonte di dati per effettuare delle operazioni di ricerca o inferenza, non risultano ad oggi affidabili quanto i sistemi chiusi e richiedono ancora molti interventi umani. Quindi sul web perdiamo molte delle caratteristiche di gestione che si potrebbero avere su dei dati puri.

Il web semantico inizia qui portando l'idea che il web non sia fatto solo di informazioni ma anche di dati in un formato adatto alla interrogazione, interpretazione e, più in generale, elaborazione automatica.

Le potenzialità date dal web semantico sono quelle che potremmo ottenere in dei sistemi chiusi.

Per esempio sarà possibile comparare i prezzi di un prodotto su scala planetaria e da un'unica interfaccia di visualizzazione/navigazione. Quindi arrivare ad un web semantico vuol dire pensare tutta una nuova serie di applicazioni che utilizzino il web come fonte dati.

Se questo dovesse avvenire ci troveremo di fronte ad una nuova rivoluzione dove la progettazione del singolo sito dovrà essere effettuata anche in visione di una scala più globale.

Per fare ciò il web semantico si avvale di un'architettura che permetta di marcare i dati e le relazioni tra essi tramite delle ontologie.

Questa architettura si basa su una serie di linguaggi per rappresentare dei grafi tramite il meccanismo delle triple (soggetto [risorsa], predicato [ontologia], oggetto [risorsa]).

Esistono almeno tre modi per descrivere tale grafo di cui la principale è RDF (Resource Description Framework) che è un linguaggio basato su XML definito dal W3C:

Il web semantico sarà certamente il futuro del web ma c'è un problema di fondo per poterlo applicare. Il problema è dato dal fatto che allo stato attuale l'espressione di grafi deve essere effettuata manualmente e in dei file esterni alle nostre pagine HTML o XHTML.

Questo (ri)porta grandi problemi di gestione e manutenzione delle strutture dati che spesso si ritrovano disallineate ma soprattutto alla scarsa possibilità che chi produce pagine web ne curi tali aspetti. Per ovviare a questo problema all'interno del W3C si stanno sviluppando una serie di soluzioni/proposte per fondere l'espressione di triple all'interno delle nostre pagine HTML o XHTML. Il W3C ha istituito una task force chiamata "RDF in XHTML" che collaborando con l'HTML e l'RDF Working Group ha lo scopo di vagliare le varie soluzioni per effettuare la fusione dei due linguaggi specialmente su XHTML 2.0.

Le soluzioni studiate da questo gruppo si dividono tra quelle per XHTML 1.0 e quelle per XHTML 2.0 dove troviamo un modulo nativo per RDF.

In questa nuova visione di web è compito degli ingegneri del software studiare le architetture ed i linguaggi ma ancora più importante è che i progettisti di interfacce o più in generale di applicativi web inizino ad immaginare cosa potranno proporre ai loro utenti in futuro potendo finalmente rispondere concretamente alle loro esigenze.