

*Termine utilizzato per descrivere il processo di progettazione, implementazione e valutazione di **spazi informativi** che siano psicologicamente e sociologicamente accettabili dagli stakeholders. Dillon (2002)*

*Definizione della **struttura** di un sistema, del modo in cui l'informazione è **raggruppata**, i metodi di **navigazione** e la **terminologia** usata entro il sistema. Il processo di costruzione delle **modalità di accesso all'informazione** finalizzato a permettere agli utenti di navigare velocemente e produttivamente all'interno del sito basandosi solamente sul loro intuito. McCracken (2005)*

Rosenfeld and Morville (2007), fornisce quattro definizioni di architettura dell'informazione:

- La combinazione di organizzazione, etichettatura e schemi di navigazione all'interno di un sistema informativo.
- La progettazione strutturale di uno spazio informativo, finalizzata a facilitare il completamento di compiti e l'accesso intuitivo ai contenuti.
- L'arte e la scienza di strutturare e classificare siti web ed intranet per aiutare gli utenti a trovare e utilizzare l'informazione.
- Una disciplina emergente, una comunità di pratiche finalizzata a portare i principi della progettazione e dell'architettura nel panorama digitale.

Il valore di un sistema informativo è legato ad una serie di aspetti

- Utilità;
- Usabilità;
- Piacevolezza;
- Reperibilità (trovabilità);
- Accessibilità;
- Credibilità;
- Valore.

<http://semanticstudios.com/publications/semantics/000029.php>

Un sistema informativo deve essere utile; deve contenere informazioni di qualità e fornire servizi utili. L'utilità di un sito web è data dalla sua capacità di fornire le informazioni che gli utenti cercano, oppure di permettere loro di portare a termine i compiti che si sono prefissi. All'interno di un sistema informativo l'utilità è spesso rappresentabile in termini di **rapporto tra i benefici ottenuti e i costi**, valutati in termini di tempo ed energie impiegate.

Secondo lo standard ISO 9241-11 l'usabilità è "Il livello in cui un prodotto può essere usato da specifici utenti per raggiungere specifici obiettivi con efficacia, efficienza e soddisfazione, in uno specifico contesto d'uso."

L'usabilità di un sistema informativo si deve occupare di due aspetti, legati ma distinti:

- l'architettura informativa.
- l'interazione;

A parità di utilità ed usabilità un sistema informativo è migliore se è piacevole da usare Buseti et al. (2003). Nel disegno industriale la bellezza non è considerata come qualcosa di astratto, ma come il risultato logico della strutturazione della forma legata alla funzione dell'artefatto stesso. Una buona progettazione estetica è fondamentalmente una questione logica, ancora prima di essere una questione psicologica, etnografica, sociologica.

È importante che le informazioni, i servizi ed i prodotti di un sistema informativo possano essere trovati facilmente dagli utenti. La reperibilità (findability – trovabilità) è uno dei compiti più importanti dell'architettura dell'informazione.

L'usabilità e gli aspetti di interaction design di un sistema informativo sono spesso indipendenti dalle dimensioni della base di dati. Nei sistemi di grandi dimensioni, però, trovare l'informazione che si cerca può essere un problema, e dunque è centrale garantirne la trovabilità.

Findability is about designing systems that help people find what they need.

Peter Morville

L'accessibilità implica il rispetto di standard finalizzati all'accesso all'informazione da parte di individui con differenti abilità, strumenti e preferenze, in molteplici contesti d'uso Lazar et al. (2004).

Web accessibility means that people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web. Web accessibility also benefits others, including older people with changing abilities due to aging.

<http://www.w3.org/WAI/intro/accessibility.php>

La credibilità è un aspetto molto importante, anche se spesso trascurato, di ogni organizzazione Fogg et al. (2001). Questo vale a maggior ragione per il web, in quanto l'assenza di contatto fisico fra l'organizzazione e l'utente rende quest'ultimo più diffidente.

È importante che il sito offra risorse o servizi di valore per gli utenti e che crei altresì valore anche per il committente Conci (2006).

- Aumentare l'**utilità** del sito in quanto alcuni dei suoi metodi empirici permettono di identificare gli interessi, le aspettative e le esigenze degli utenti.
- Migliorare l'**usabilità** in quanto permette di rilevare il lessico degli utenti e di cogliere i loro modelli mentali impliciti concernenti il modo in cui si aspettano che l'informazione sia strutturata e categorizzata.
- Migliorare la **reperibilità** di un sistema informativo. Per essere reperibile l'informazione deve infatti essere strutturata in maniera coerente, in modo da rispettare le aspettative implicite degli utenti.

- l'**identificazione dei contenuti** che gli utenti si aspettano di trovare in un sito web;
- la valutazione dell'**importanza** che gli utenti attribuiscono ai contenuti;
- la conoscenza del **lessico** adottato dagli utenti, ed il conseguente adattamento della terminologia del sito;
- la **strutturazione** delle unità informative in partizioni gerarchiche (alberi);
- l'identificazione delle risorse la cui collocazione può risultare problematica;
- la progettazione di **metainformazioni** sulle risorse.

<http://www.iawiki.net/DefiningTheDamnThing>

Vi sono tre dimensioni che vanno tenute in considerazione nella progettazione di un sistema informativo Rosenfeld and Morville (2007):

- il contesto: gli scopi del committente, le politiche, la cultura, la tecnologia, le risorse, i vincoli;
- i contenuti del sistema informativo: i documenti, i file, le applicazioni, i servizi, i metadati;
- gli utenti del sistema.

Lo standard ISO 13407 definisce lo user centred design come “Un approccio allo sviluppo di sistemi interattivi focalizzato specificatamente all’usabilità. È una attività multidisciplinare, che richiede competenze e tecniche specifiche di ergonomia . . . Applicare l’ergonomia al disegno di sistemi richiede di considerare fattori primari le capacità, competenze, conoscenze, limitazioni ed esigenze degli utenti.”

Si assume che nessuno conosca competenze, cultura, bisogni, limiti, atteggiamenti degli utenti reali meglio degli utenti medesimi, e pertanto prevede il **coinvolgimento degli utenti** in tutte le fasi della progettazione, realizzazione e gestione di un prodotto.

La metodologia centrata sull'utente prevede il coinvolgimento attivo degli utenti, la comprensione dei requisiti degli utenti e dei compiti, l'allocazione appropriata di funzioni tra gli utenti e il sistema, un approccio iterativo alla progettazione Mao et al. (2001).

Coinvolgere gli utenti nel design di un sistema informativo ci aiuta a capire:

- Cosa si aspettano di trovare, gli utenti, nel sito che stiamo costruendo?
- Come si aspettano che l'informazione sia strutturata, organizzata, classificata e presentata?
- Qual'è il loro lessico? Quali termini dobbiamo usare per aiutarli a comprendere, identificare e recuperare i contenuti?

Rosenfeld and Morville (2007) identificano tre aree di intervento:

- **Organizzazione:** il modo in cui l'informazione deve essere organizzata e strutturata.
- **Labeling:** il lessico da usare per etichettare le risorse e le categorie del sistema di informazione.
- **Navigazione e ricerca:** gli strumenti dell'interfaccia che permettono agli utenti di navigare nel sistema e di cercare le informazioni attraverso il motore di ricerca.

Gli schemi esatti sono quelli che si basano su una classificazione od un ordinamento esatto e non ambiguo. Gli esempi più importanti:

- ordine alfabetico;
- ordine cronologico;
- classificazione geografica.

Generalmente l'ordinamento può essere fatto automaticamente dall'applicazione che presenta i dati.

Gli schemi ambigui sono meno semplici da implementare, ed introducono degli elementi di soggettività. Ciononostante risultano spesso estremamente utili. Esempi

- per argomento;
- per compito;
- per tipologia di audience;
- in base a metafore;
- schemi ibridi.

È proprio nella creazione di questa tipologia di schemi che diviene necessario, in fase di design, coinvolgere gli utenti utilizzando degli strumenti di elicitazione della conoscenza.

Le strutture organizzative si occupano della modalità di immagazzinamento e presentazione delle informazioni.

L'adozione di appropriate strutture organizzative è soggetta a tre fattori: il fattore tecnologico, la tipologia dei contenuti, gli aspetti di usabilità legati alla modalità di presentazione dei dati.

Rosenfeld and Morville (2007) identificano tre tipi di strutture:

gerarchiche, **tabellari** ed **ipertestuali**. A nostro avviso altri due tipi di strutture meritano la nostra attenzione, poiché sono delle tipologie di organizzazione dell'informazione emergenti in internet: la **classificazione a faccette** e la **classificazione a parole chiave (tags)**.

Una struttura gerarchica divide il dominio semantico della struttura informativa in partizioni. Un tipico esempio – di struttura gerarchica esatta – è la tassonomia linneiana degli esseri viventi.

Su internet vi sono degli esempi estremamente celebri, le directory come dmoz.org o yahoo. In questo caso, però, la struttura gerarchica è di tipo ambiguo.

Strutture a tabella, o database

Esistono numerosi esempi di questa struttura: il client di posta elettronica elenca la posta in arrivo in una tabella, dove ad ogni riga (record) corrisponde una mail, mentre le colonne rappresentano i diversi attributi delle mail: data, mittente, oggetto, dimensione in kilobites.

Un altro esempio è costituito dal servizio offerto da *librarything.com*, un sito permette di condividere la propria libreria, inserendo la lista dei libri che si possiede. Chi entra nella mia *libreria virtuale* (cliccando su www.librarything.com) troverà la lista dei miei libri, che potrà essere ordinata per autore, per titolo, per punteggio di gradimento.

Struttura ipertestuale – *network*

La struttura ipertestuale costituisce la più importante caratteristica distintiva del web. Il web è, dal punto di vista dell'utente, una ragnatela di risorse testuali o multimediali fra loro collegati attraverso dei collegamenti ipertestuali.

Un sito web non è un sito web se non ha collegamenti ipertestuali, in quanto questi costituiscono lo strumento universale di navigazione all'interno del web. E dunque anche le strutture gerarchiche e tabellari usano i collegamenti ipertestuali per la navigazione. Vi sono però dei sistemi informativi dove i collegamenti ipertestuali costituiscono e rappresentano la struttura. L'esempio più importante è costituito da wikipedia ¹. Questo tipo di struttura è particolarmente adatto per le conoscenze di tipo enciclopedico.

¹it.wikipedia.org nella versione italiana

In termini molto sintetici la classificazione a faccette è una classificazione multidimensionale.

Alcuni importanti siti di commercio elettronico fanno uso di questo tipo di classificazione. Se cerchiamo una macchina fotografica digitale su siti come froogle o ebay ci viene offerta la possibilità di restringere la ricerca in base a differenti parametri: tipo di fotocamera (compatta, reflex), risoluzione in megapixels, zoom, marca, fascia di prezzo.

Dal punto di vista dell'interazione con l'utente la classificazione a faccette si propone di combinare la strutturazione della classificazione gerarchica con la multidimensionalità delle strutture a tabella.

I servizi che si basano su questa struttura permettono agli editori delle informazioni, ma anche agli utenti, di aggiungere delle informazioni alle risorse (documenti, siti web, immagini, video). Queste informazioni, che tecnicamente sono dei metadati, sono generalmente delle parole chiave, tag in inglese. I siti web che si basano su queste strutture permettono agli utenti di navigare all'interno del sito proprio attraverso le parole chiave.

Flickr ² è un sito che permette di condividere le proprie fotografie. L'utente si registra e carica sul server di flickr le proprie immagini digitali. Il sito invita a contrassegnare ogni immagine con una o più parole chiave. Gli utenti potranno visualizzare le fotografie presenti sul sito attraverso la ricerca per parola chiave. del.icio.us ³ offre un servizio di bookmark online. Se, navigando, mi imbatto in un sito o una pagina che reputo interessante, posso decidere di salvarne il collegamento fra i segnalibri del mio browser. Del.icio.us permette di fare la stessa cosa salvando l'indirizzo sul loro sito, e contrassegnandolo con dei tag.

²www.flickr.com

³del.icio.us

Con l'avvento di servizi come Google Maps ⁴ è possibile associare ad una risorsa delle coordinate geografiche. Diviene così possibile navigare le risorse attraverso una mappa satellitare. Flickr offre un servizio di questo genere, che permette di visualizzare il luogo dove sono state scattate le fotografie.

⁴maps.google.com

Secondo Pirolli and Card (1999) gli esseri umani sono degli *informavori* il cui successo adattativo dipende dalla loro capacità di applicare con successo delle sofisticate strategie di selezione delle informazioni, di attribuzione di senso, di problem solving e decision making.

Secondo la *Information Foraging theory* da loro proposta la ricerca e la selezione delle informazioni può essere paragonata alla strategia di foraggiamento degli animali, e dunque i sistemi informativi dovrebbero massimizzare il rapporto fra il valore dell'informazione per l'individuo ed il costo sostenuto per trovarla, analizzarla, elaborarla.

Nel contesto dei siti internet le risorse sono presentate all'utente attraverso dei link testuali o grafici.

Il designer utilizza le etichette come dei suggerimenti prossimali finalizzati a permettere all'utente di intuire i contenuti del documento collegato; nella metafora del foraggiamento informativo questi suggerimenti vengono definiti da Chi et al. (2000) *information scent*: ciò che il link suggerisce è una percezione soggettiva del valore, del costo e delle modalità di accesso alle informazioni. L'utente è guidato nella sua esplorazione della struttura informativa da questi suggerimenti Chi et al. (2001).

Con etichetta si intende un'unità informativa di piccole dimensioni, finalizzata ad identificare una risorsa: un documento, un file audio, un prodotto, un filmato, una persona. L'etichetta è generalmente testuale, ma può essere anche una piccola immagine (un'icona) e, in determinati casi, anche un suono (es. una suoneria personalizzata nel cellulare che ci dice chi ci sta chiamando).

Vanno dunque usate delle etichette che sappiano guidare in maniera corretta gli utenti, permettendo loro di intuire la natura dell'informazione a cui sono collegati.

Nella costruzione di un sistema informativo la progettazione assorbe, generalmente, tempo e risorse, tanto che spesso i clienti – ma anche alcuni sviluppatori – sono tentati di saltare intere fasi progettuali. In realtà una buona progettazione permette di creare siti più utili ed usabili; se la progettazione è orientata non solo al presente ma anche agli sviluppi futuri sarà meno soggetta ad obsolescenza e più facile da aggiornare Fuccella and Pizzolato (1998). Caprio and Ghiglione (2003) identificano quattro fasi di progettazione

- **discovery**: identificazione degli obiettivi del sito, definizione dei requisiti, evidenziazione di eventuali vincoli progettuali;
- **analisi**: prevalentemente l'inventario dei contenuti e l'analisi dei profili utente;
- **architettura**: labeling, categorizzazione dei contenuti, navigazione, definizione del database;
- **sviluppo**: il sito viene sviluppato e testato.

Il processo che presentiamo costituisce un percorso ideale, che integra differenti proposte Fuccella and Pizzolato (1998); Caprio and Ghiglione (2003); Sinha and Boutelle (2004); McGovern (2002); McQuaid et al. (2003).

Un sito internet deve produrre valore per chi lo commissiona, in modo che vi sia un ritorno degli investimenti. Il ritorno degli investimenti di un'azienda può essere definito in termini di differenti variabili Conci (2006): aumento della produttività dei dipendenti, diminuzione dei costi di formazione, allungamento dei cicli di vita del sito Fuccella and Pizzolato (1998), aumento delle vendite (on line e off line), aumento della notorietà del marchio Zeni (2006), diminuzione dell'uso del call center da parte dei clienti.

Risulta pertanto necessario capire quali sono gli obiettivi dell'azienda committente. Questo passaggio rientra nella fase di stakeholder analysis Sinha and Boutelle (2004).

Caprio and Ghiglione (2003) suggeriscono di adottare il metodo dell'intervista semistruutturata, da sottoporre a tutti gli stakeholder.

Nell'intervista vanno chiesti:

- gli obiettivi dell'azienda (l'obiettivo principale, gli obiettivi a breve, medio e lungo termine, in ordine di importanza);
- le motivazioni, le aspettative in merito al sito web;
- il target di utenza a cui pensano il sito debba rivolgersi;
- i criteri di successo del sito.

Uno degli svantaggi di una progettazione esclusivamente normativa è che tende ad assumere di conoscere gli utenti e le loro caratteristiche. Questa assunzione si rivela però spesso errata Nielsen (1996). Conoscere il profilo degli utenti è molto importante nella costruzione di un sistema informativo che intenda soddisfare le loro esigenze. Nell'identificazione di tale profilo è importante trovare un metodo di campionamento dei partecipanti che ne selezioni un gruppo rappresentativo. Fuccella and Pizzolato (1998) indicano, come possibile fonte, i dati di una analisi di marketing. Questa fonte però non sempre è adeguata. In primo luogo perché non tutte le organizzazioni dispongono di analisi di questo genere. In seconda istanza non è detto che l'utenza del sito internet sia sovrapponibile a quella emersa dall'analisi di marketing.

Fuccella and Pizzolato (1998) distinguono fra *passive* e *active survey collection*: nella ricerca attiva il designer va a caccia di partecipanti, attraverso una campagna pubblicitaria, o utilizzando una mailing list o un gruppo di discussione. La *passive collection* consiste nell'utilizzare il sito internet esistente nella raccolta di partecipanti: all'interno del sito viene presentato un invito a partecipare al questionario. Questo è, a nostro avviso, il metodo migliore, in quanto ci assicura il miglior campionamento: i partecipanti che rispondono sono i reali utilizzatori del sito. Come vedremo nelle sezioni successive l'uso di strumenti web per la somministrazione dei questionari quali il free listing, la valutazione di importanza e il card sorting sono motivati anche dalla possibilità di testare, on site (nel senso letterale del termine) i reali utenti del sito.

Lo svantaggio di questo metodo, però, è che esclude dall'analisi potenziali nuovi utenti; vi è inoltre il rischio che alcune categorie di utenti siano più motivati di altri a rispondere, portando a veri e propri errori di campionamento. Risulterebbe dunque molto utile poter disporre di differenti modalità di reclutamento, e poter distinguere i partecipanti in base alla modalità, al fine di valutare se i risultati che si ottengono sono significativamente diversi. In ogni caso le possibili difficoltà non debbono indurre i progettisti a rinunciare. Poiché la finalità è applicativa (e non scientifica) un campionamento *sbilanciato* è pur sempre meglio di nulla.

Fuccella and Pizzolato (1998) suggeriscono la somministrazione di brevi questionari finalizzati a delineare alcuni profili di base degli utenti:

- profilo anagrafico: sesso, età;
- profilo professionale: titolo di studio, professione;
- profilo di utilizzo del web: come, quando, perché usa internet;
- se l'utente è stato contattato attraverso il sito esistente, possono venir chieste anche delle informazioni sull'uso del sito, su pregi e difetti identificati o desiderata.

Queste informazioni, comunque, possono essere raccolte anche in fase di somministrazione di questionari più specifici, come il free listing ed il card sorting.

I questionari possono fornire delle utili informazioni sugli utenti e sulle loro richieste. Va però tenuto conto che non sempre gli utenti sono in grado di dire cosa vogliono o cosa sia meglio per loro. È pertanto necessario verificare non solo le opinioni ma l'uso reale, ad esempio attraverso l'analisi contestuale.

Con alcuni degli utenti è possibile realizzare delle interviste, finalizzate a comprendere i loro comportamenti, bisogni e aspettative Caprio and Ghiglione (2003). Dalle interviste e dai questionari è possibile delineare dei profili utente. Alcuni autori suggeriscono di utilizzare i profili più rappresentativi per creare delle *personas*, dei personaggi fittizi e verosimili su cui focalizzarsi nel design del sito web Sinha (2003b).

Il vantaggio delle interviste è che permettono di approfondire la conoscenza di alcuni profili di utenti. Lo svantaggio principale è che è costosa, e dunque può essere somministrata ad un numero limitato di persone.

L'indagine contestuale consiste nell'osservare l'utente durante la sua attività e nel luogo in cui si svolge Caprio and Ghiglione (2003). Questo strumento permette di cogliere le esigenze ed i comportamenti degli utenti. Un approccio di questo genere permette di rendere esplicite conoscenze o esigenze tacite, di cui l'utente non è consapevole ma che di fatto ne condizionano l'interazione con il sistema informativo.

McQuaid et al. (2003), nel ridefinire l'architettura dell'informazione di una biblioteca pubblica, decisero di "walk a mile in the customers' shoes":

Dopo aver osservato gli utenti e parlato con i bibliotecari, avevamo un'idea molto più completa del tipo di informazioni disponibili e del modo in cui le persone vi accedono. Abbiamo scoperto, ad esempio, che l'informazione che un utente sta cercando può risiedere in media diversi (libri, bollettini, riviste, microfiche, giornali, videocassette, poster, articoli elettronici, ed altre persone) in luoghi diversi con metodi di accesso ed organizzazione diversi (sistema Dewey, Library of Congress, collezioni speciali). McQuaid et al. (2003)

Se ci si sta occupando del redesign di un sito esistente, il primo passo consiste nell'analisi dei contenuti del vecchio sito. In primo luogo va creata una lista delle risorse presenti Caprio and Ghiglione (2003). La lista può essere integrata con altre informazioni legate all'uso del sito Gamberini and Valentini (2001):

- il numero medio di contatti giornalieri per ogni pagina, basandosi sul file di log del server;
- i referrer alla pagina, ovvero l'elenco di pagine di altri siti web che hanno un link a quella pagina;
- la visibilità della pagina sui motori di ricerca, e le parole chiave che indirizzano i motori a quella pagina;
- eventuali commenti o voti alla pagina, se il cms lo permette.

Per identificare il dominio semantico è utile elencare le risorse informative presenti su siti internet concorrenti. Questo metodo è particolarmente utile se si sta costruendo un sito ex novo, ma può dare indicazioni utili anche nel caso di redesign di un sito esistente Cordioli (2006).

L'analisi competitiva può essere finalizzata non solo a definire il dominio semantico, ma anche ad identificare eventuali pratiche virtuose, standard e consuetudini nel segmento di mercato considerato.

Caprio and Ghiglione (2003) sottolineano come da un'analisi competitiva sia possibile cogliere diversi aspetti dei siti concorrenti:

- Caratteristiche generali: impressioni, categoria del sito, profilo del target, stile del sito.
- Struttura: aree generali, aree specifiche per profili utenti, organizzazione dei contenuti, navigazione.
- Funzionalità: motori di ricerca, help, registrazione, autenticazione.

In questa fase di processo l'analisi è finalizzata ad ottenere una lista delle risorse presenti sul sito concorrente, in maniera simile alla lista delle risorse esistenti.

Nel definire un dominio semantico ci si pone due domande principali. La prima domanda è: “quali sono i contenuti del dominio?” La seconda domanda è: “come sono strutturati i contenuti?”. Il free listing è una tecnica che può aiutarci a determinare l'ampiezza del dominio e fornire alcune intuizioni su come il dominio è strutturato. – Sinha (2003a)

La tecnica del free listing può essere utilizzata per coinvolgere gli utenti nella definizione dei contenuti Coxon (1999). Più in particolare può essere usata per due funzioni: elencare i contenuti, l'ambito e i confini del dominio semantico; identificare il lessico degli utenti.

Dopo aver utilizzato alcuni o tutti i metodi elencati sarà necessario elencare le voci così ottenute in un'unica lista, badando naturalmente ad eliminare le ripetizioni e le ridondanze.

È importante includere in questa lista finale tutte le voci, e non solo quelle corrispondenti a risorse già implementate nel sito web; attraverso la valutazione di importanza sarà possibile identificare le aree informative sulle quali varrà la pena di focalizzarsi.

La valutazione dell'importanza delle risorse consiste in un questionario in cui vengono elencate le risorse identificate nella fase precedente e viene chiesto ai partecipanti di esprimere, attraverso una scala Likert, quanto ritengano importante ognuna delle voci elencate Rugg and McGeorge (1997).

Da questo questionario si otterrà una classifica dell'importanza attribuita dagli utenti alle risorse.

Può essere utile analizzare separatamente i risultati del questionario per segmenti di utenti diversi.

Valutazione di importanza: scopi

- Permettere agli editori del sito di identificare gli argomenti sui quali è importante concentrare l'attenzione nella fase di sviluppo e aggiornamento dei contenuti.
- Decidere a quali risorse dare maggiore risalto nel sito internet, magari attraverso dei link nella home page.
- Se nella definizione dell'utenza sono emersi gruppi differenti, è possibile che i diversi gruppi attribuiscono un'importanza diversa a risorse differenti; attraverso il questionario è possibile far emergere queste differenze, delle quali è necessario tener conto nella progettazione della navigazione.
- Nel card sorting, agli utenti si chiede di classificare una lista di elementi; la prestazione ottimale dei partecipanti si ha quando la lista non supera i 60 - 70 elementi. Se l'elenco di cui disponiamo è più lungo può essere opportuno sottoporre a card sorting solo le 60 voci considerate più importanti dagli utenti. Successivamente è possibile somministrare un secondo card sorting con le voci escluse.

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 relative alla categorizzazione dei contenuti di un sito web Nielsen and Sano (1994); Rugg and McGeorge (1997); Maurer and Warfel (2004); Nielsen (2004); Fincher and Tenenberg (2005).

Nell'ambito dell'architettura dell'informazione il card sorting costituisce un metodo di design centrato sull'utente, finalizzato ad ottimizzare la reperibilità (*findability*) di un sistema.

Il card sorting costituisce un metodo efficace per individuare i modelli mentali impliciti degli utenti, rendendo esplicite le loro aspettative di categorizzazione dei contenuti. Conoscere i modelli mentali e le categorizzazioni implicite ci permette di organizzare le informazioni in modo che siano più facili da trovare e da utilizzare, migliorando la qualità del prodotto.

Attraverso il card sorting è possibile identificare il criterio di classificazione usato dagli utenti ed identificare il contenuto e l'etichetta delle categorie da essi utilizzati. È possibile far emergere eventuali differenze nella categorizzazione fra diversi gruppi di partecipanti.

Il card sorting è usato da decenni nelle scienze sociali come strumento per classificare oggetti in categorie. È la tecnica di elicitazione della conoscenza più usata nell'area dell'interazione uomo-computer per far emergere i modelli mentali degli utenti, rendendo esplicite le loro aspettative di categorizzazione dei contenuti Nielsen and Sano (1994); Rugg and McGeorge (1997); Maurer and Warfel (2004); Fincher and Tenenberg (2005); costituisce un metodo di design centrato sull'utente, finalizzato ad ottimizzare la reperibilità (findability) all'interno un sistema. Se i partecipanti sono rappresentativi degli utenti del sito i risultati dell'analisi tenderanno a riflettere la struttura in cui gli utenti si aspettano che le informazioni siano presentate. È dunque un buon punto di partenza per organizzare la struttura del sito web.

Secondo Rugg and McGeorge (1997) il card sorting può essere usato sia come tecnica esplorativa che come tecnica di classificazione vera e propria. Secondo questi autori il card sorting può essere applicato ad una gamma di entità estremamente ampia, che spazia da elenchi di oggetti concreti a concetti astratti, e può essere utilizzata ricorsivamente a vari livelli di una struttura informativa. Attraverso il card sorting possiamo far emergere:

- i criteri che i partecipanti adottano per categorizzare e cercare le informazioni
- la struttura informativa che implicitamente si aspettano di trovare
- le eventuali differenze fra diversi gruppi di utenti
- le etichette delle categorie, espresse nel lessico degli utenti.

È una tecnica facile da realizzare e facile da far comprendere ai partecipanti: gli utenti la considerano un metodo di classificazione naturale.

Le circostanze ideali per ottenere dei buoni risultati sono:

- l'elenco non è superiore a 60-70 elementi
- i contenuti sono fra loro omogenei
- i partecipanti conoscono e comprendono i contenuti.

Card sorting aperto e chiuso

Il card sorting può essere somministrato in due modalità:

- card sorting aperto
- 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 card sorting aperto sono particolarmente interessanti poiché possono darci informazioni su delle tipologie di categorizzazione non immaginate a priori dallo sperimentatore, anche se la maggiore libertà concessa all'utente rende meno coerenti i risultati raccolti.

Analisi delle componenti principali

L'analisi delle componenti principali (PCA) è una tecnica statistica esplorativa multivariata finalizzata a semplificare insiemi di dati complessi Anderson and Gerbing (1988); Raychaudhuri et al. (2000); Ding and He (2004). Date m osservazioni su n variabili, lo scopo della PCA è di ridurre la dimensionalità della matrice di dati trovando r nuove variabili, dove $r < n$. Queste r variabili, definite componenti principali, hanno la proprietà di *spiegare* la varianza delle n variabili originali e di essere fra loro ortogonali e non correlate. Ogni componente principale è una combinazione lineare delle variabili originarie, ed analizzando i coefficienti è possibile attribuire un significato alle componenti Raychaudhuri et al. (2000). Nelle nostre analisi (sulle matrici di prossimità) i risultati ottenuti con lo scaling multidimensionale classico (metrico) e la PCA sono del tutto equivalenti.

La clusterizzazione è una divisione di un insieme in gruppi di oggetti fra loro simili. Ogni gruppo, definito cluster, consiste di oggetti che sono simili fra di loro e dissimili dagli oggetti degli altri gruppi Berkhin (2007). La cluster analysis è un sistema di classificazione esplorativo senza supervisione Xu and Wunsch (2005) che costruisce una partizione, ovvero un insieme di gruppi fra loro disgiunti Ding and He (2004).

Vi sono innumerevoli algoritmi di classificazione; i più comuni si distinguono in metodi gerarchici e metodi di partizionamento Berkhin (2007). La cluster analysis gerarchica è un metodo gerarchico agglomerativo, mentre la k-means è un algoritmo di partizionamento.

La cluster analysis gerarchica costruisce un albero di clusters, detto dendrogramma o albero di classificazione gerarchica Coxon (1999); Sinha and Boutelle (2004); questo è l'algoritmo di clusterizzazione più frequentemente applicato alle matrici di prossimità e dunque al card sorting Tullis and Wood (2004); Faiks and Hyland (2000); Berkhin (2007).

K-means è un algoritmo di partizionamento che assegna l'insieme di oggetti in K clusters Xu and Wunsch (2005); Berkhin (2007); vantaggi:

- permette di applicare ai dati del card sorting non solo l'algoritmo di clusterizzazione gerarchica, ma anche un algoritmo di partizionamento;
- permette di visualizzare i risultati in uno spazio bidimensionale (grazie alla PCA);
- permette di far emergere delle dimensioni semanticamente interpretabili; poiché il clustering si basa su tali dimensioni, è possibile interpretare la classificazione in base a tali variabili *latenti* Raychaudhuri et al. (2000).

In The elements of user experience, lo skeleton è il 4' livello. A questo livello, secondo Cooper, vanno definiti:

- la piattaforma (web, mobile ...)
- gli elementi funzionali e i dati da rappresentare
- il raggruppamento e le gerarchie
- il layout visivo.

Come vedremo nella lezione dedicata all'usabilità, è necessario rispettare euristiche e linee guida, ed è opportuno testare frequentemente questi passaggi.

Riprendendo la slide dell'introduzione, JJG definisce 3 componenti:

- l'**information design**: la presentazione delle informazioni all'utente;
- l'**interface design**: la progettazione degli elementi dell'interfaccia per permettere agli utenti di interagire con l'applicazione;
- la progettazione della **navigazione**, che permette agli utenti di muoversi all'interno della struttura informativa.

Uso di convenzioni, metafore, pattern, linee guida. Vengono prodotti wireframes.

Dalla macrostruttura (sito) alla microstruttura (pagina)

Gli strumenti adottati fino a questo punto sono finalizzati a definire la struttura del sito.

A questo punto si ragiona sulle pagine. Vanno definiti:

- le varie tipologie di pagine (home, indice di sezione, pagina documento, pagina prodotto ...)
- per ogni tipologia, un template (wireframe) in cui disporre gli elementi informativi e funzionali.

Dal punto di vista implementativo (back-end) gli sviluppatori devono definire la struttura del database che gestirà le informazioni.

Ad esempio, per un sito che vende libri, dovremo immaginare (almeno) 2 tabelle:

- la tabella **autori** (nome, cognome, data di nascita, nazionalità, biografia);
- la tabella **libri** (idAutore, titolo, anno di pubblicazione, casa editrice ...).

Sul lato dell'interfaccia, dobbiamo creare lo scheletro per la pagina autore e quello per la pagina libro, decidendo quali informazioni presentare, e come.

Vanno definiti:

- l'importanza relativa degli elementi, e la loro gerarchia;
- la disposizione nello spazio, tenendo conto di:
 - gerarchia e importanza;
 - eventuali sequenze logiche;
 - raggruppamenti funzionali e informativi

I wireframes sono finalizzati a creare dei prototipi delle pagine, basati sui criteri sopra esposti.

Per creare wireframes possiamo usare strumenti diversi:

- carta e penna
- software dedicati
- powerpoint, inkscape
- direttamente in html

È buona norma iniziare con carta e penna, in modo da produrre molte idee in poco tempo e con poco sforzo. Wireframes troppo dettagliati sono controproducenti, perché si rischia di affezionarsi ad una ipotesi ancora embrionale.

- Yahoo design patterns
- Welie patterns
- Patternry
- ui-patterns
- time-tripper

Esercizio: identificare il layout di questi siti

yahoo.com, google.com, youtube.com, live.com, msn.com, myspace.com, wikipedia.org, facebook.com, blogger.com, orkut.com, microsoft.com, ebay.it, aol.com, amazon.com, The Internet Movie Database (imdb.com), wordpress.com, flickr.com, bbc.co.uk, craigslist.org, cnn.com

- permettere all'utente di capire dov'è
- permettere all'utente di muoversi all'interno del sito
- permettergli di capire il contesto e le modalità di organizzazione delle informazioni
- dare un'idea di quali informazioni può trovare (information scent)

- navigazione **globale**: fornisce l'accesso ai nodi principali dell'intero sito. Generalmente è presente in ogni pagina
- navigazione **locale**: fornisce l'accesso alle voci appartenenti alla stessa categoria della pagina
- navigazione **contestuale**: accesso a voci semanticamente correlate
- navigazione **supplementare**: accesso ad altre voci, importanti o obbligatorie (legalese)
- navigazione **di cortesia**: link a pagina indice, mappa del sito ...
- **briciole di pane**: *home > categoria1 > categoria1.A > pagina*

Esercizio: identificare i sistemi di navigazione

Identificare la navigazione globale, locale, contestuale, supplementare, di cortesia, briciole di pane.

yahoo.com, google.com, youtube.com, live.com, msn.com, myspace.com, wikipedia.org, facebook.com, blogger.com, orkut.com, microsoft.com, ebay.it, aol.com, amazon.com, The Internet Movie Database (imdb.com), wordpress.com, flickr.com, bbc.co.uk, craigslist.org, cnn.com

Grazie

stefano@bussolon.it

<http://www.bussolon.it>

- Anderson, J. C. and Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3):411–423.
- Berkhin, P. (2007). Survey of clustering data mining techniques. [Online; accessed 29-August-2007].
- Busetti, T., Bussolon, S., and Sartori, G. (2003). Il colore dei link delle pagine web come memoria esterna. In *Atti del Congresso nazionale dell'Associazione Italiana di Psicologia, Sezione di Psicologia Sperimentale*, Bari (I).
- Caprio, L. and Ghiglione, B. (2003). *Information Architecture*. Tecniche Nuove, Milano.
- Chi, E., Pirolli, P., and Pitkow, J. (2000). The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a web site. In *CHI Letters*, volume 2. CHI.
- Chi, E. H., Pirolli, P., Chen, K., and Pitkow, J. (2001). Using information scent to model user information needs and actions on the web.
- Conci, M. (2006). Portale unitn: architettura dell'informazione centrata