POLITECNICO DI MILANO
Facolta di Ingegneria dell'Informazione

POLO REGIONALE DI COMO
Corso di Laurea Specialistica in Ingegneria Informatica

INTEGRATING SEARCH AND INFORMATION ARCHITECTURE BEYOUND IMPROVING NAVIGATIONAL PATTERN

Supervisor: **Prof. PAOLO PAOLINI**

Master Graduation Thesis by: **FATEMEH RAZZAGHI**

Student Id. number: **720346**

Academic Year 2010/2011

POLITECNICO DI MILANO
Facolta di Ingegneria dell'Informazione

POLO REGIONALE DI COMO
Corso di Laurea Specialistica in Ingegneria Informatica

INTEGRATING SEARCH AND INFORMATION ARCHITECTURE BEYOUND IMPROVING NAVIGATIONAL PATTERN

Relatore: **Prof. PAOLO PAOLINI**

Tesi di laurea di: **FATEMEH RAZZAGHI**

matr. **720346**

Anno Accademico 2010/2011

ACKNOWLEDGMENTS

I am heartily thankful to my supervisor, Professor Paolo Paolini, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject.

Special thanks also to Luigi Spagnolo who has made available his supports in a number of ways during this work.

I would also like to convey thanks to my faculty professors and administration departments for their kind and generous support during my master program.

And last but not least, I wish to express my love and gratitude to my beloved husband and families; for their understanding & endless love, through the duration of my studies.

Table of Contents

Abstract in English language

The research in web engineering has developed among the years a consolidated set of agile design methodologies and technologies that allowed the implementation of very large and robust web applications and information systems. However, the traditional way of structuring information in communication-intensive websites (e.g. institutional and cultural heritage websites, e-commerce applications, etc.) is no more able to cope with corpuses of contents that are getting larger and larger. New design approaches are required to better support some crucial user experience requirements such as findability (the ability of users to locate the desired information) and sense making (the possibility of understating the coverage of the website and how information is interconnected).

The present thesis is aimed at illustrating the problems of traditional information architectures for interactive multimedia (web) applications and their possible solutions, as they are emerging from the recent research in the field.

Such new approaches are based on the idea of integrating the infrastructure for accessing contents – as it provided by pages and links – with a novel "search-driven" philosophy, with the purpose of enhancing both findability and sense making. Hence, the structure of a website (or of any other interactive application) becomes a seamless mix of organized information (as it has been traditionally done till now) and organized search space based on advanced navigational patterns such as faceted search and relational browsing.

With respect to the standard way of using search engines, the originality of such new design strategies lays in the fact that search is not added to the information architecture as an alternative mechanism, but it is rather part of it in an integrated manner, according to designer wishes. Integrating search and information architecture could provide several advantages to users and developers. Users have the perception of an information space that is dynamically reshaped as long as they proceed in their exploration, in a way that is adaptive to their needs. For designers and developers advantages include a more flexible management of design effort and a more efficient and effective use of content.

Abstract in lingua Italiana

Nel corso degli anni, la ricerca nell'ambito del web engineering ha sviluppato un insieme consolidato di metodologie di design e tecnologie atte all'implementazione di applicazioni web e sistemi informativi robusti e su larga scala. Tuttavia, la maniera tradizionale di strutturare le informazioni per siti communication-intensive (ossia votati alla comunicazione: ad es. siti di istituzionali o nell'ambito della comunicazione culturale, applicazione per e-commerce, ecc.) non è più sufficiente a gestire un ammontare sempre crescente di contenuti. Sono necessarie pertanto nuove strategie di design per supportare requisiti cruciali per l'esperienza utente in questo ambito di siti, quali la cosidetta findability (la facilità con cui l'utente è in grado di reparire le informazioni) e il sense making (ossia la possibilità per l'utente di comprendere l'ambito di interesse del sito e le interconnessioni esistenti nelle informazioni reperibili attraverso di esso).

Scopo della presente tesi è quello di illustrare le problematiche inerenti alle architetture dell'informazione tradizionali per applicazioni web multimediali e le relative possibili soluzioni, così come emergono dalla recente ricerca in tale campo.

Tali nuovi approcci sono basati sull'idea di integrare l'infrastruttura necessaria ad accedere ai contenuti – composta da pagine e link – con una "filosofia" della ricerca (intesa come search), allo scopo di migliorare sia la findability che il sense making. La struttura del sito (o di qualsiasi altra applicazione interattiva) diventa un insieme integrato di informazioni organizzate (così come accadeva fino ad oggi) e spazi di ricerca basati su meccanismi avanzati di navigazione come la faceted search e il relational browsing.

Rispetto al tradizionale impiego di motori di ricerca all'interno di siti web, l'originalità di questi approcci risiede nel fatto che la search non consituisce un elemento aggiuntivo e alternativo, ma diventa parte integrante del sistema, in base ai requisiti del progettista.

L'integrazione di architetture dell'informazione e ricerca è potenzialmente vantaggiosa sia per gli utenti che per designer e sviluppatori. In particolare, gli utenti possono avere la percezione di uno spazio di informazioni che si rimodella man mano che essi procedono

nell'esplorazione del sito, in maniera adattiva rispetto ai loro bisogni e desideri. Per progettisti e sviluppatori i vantaggi comprendono una maggiore flessibilità nella gestione delle attività di design, e un un più efficace ed efficiente riuso dei contenuti.

1 Introduction and Background

Humans can be defined as informavors, because they need information to survive and to accomplish their goals. In the digital age, an overwhelming amount of information is at our disposal at low costs or even for free. Ill structured, difficult to find information is, however, nearly useless.

According to Rosenfeld and Morville's work (2002)[1], which combines principles from information and library science with models from human-computer interaction and user-centered design has had great influence, and many practitioners view it as the canonical work on IA. Other discussions of IA such as Brinck, Gergle, and Wood (2002)[23] take generally similar approaches.

In these frameworks, the information architect is expected to develop an understanding of users' information needs and mental models, using research methods such as interviews, contextual inquiry, surveys, and card sorting. Based on this research, the information architect then develops an architecture that can support users' tasks and documents this architecture using standardized deliverables such as blueprints (or sitemaps) and wireframes (page diagrams or low-fidelity prototypes). The information architect then evaluates the architecture through usability testing or similar methods, and works with interaction and visual designers and developers to implement the information architecture in a functional system.

The main aim of the thesis work is to illustrate various latent and sensible problems which are perceptible in concepts of information architecture, search and navigation due to investigate and analyze different approaches in web design .The main concerns is to ingenerate these dependent concepts having dynamic navigation structures which can help large website reduces the complexity and rigidity of classical hierarchical navigation structures, without losing the overall communicative strength.

The usual approach in web design is focused on the info-architecture as main and often exclusive user interface to access contents in a website. In particular, the well-known theory of information scent states that users act like "information hunting animals": starting from the homepage, they scan for trigger words in links and then navigate through internal pages, going further or back according in their navigation to their perception of getting closer or conversely farther to the information they are looking for.

IDM (Interactive dialogue model) follows this mainstream and concerns on the metaphor of interaction as a dialogue with the web application, in which the user "talks to" the website – step by step through navigation among pages – by selecting a link to single or multiple topics from the several ones "proposed" by the current page.

Nevertheless, in the case of large websites, such the ones of public administration, governments, corporations or universities, "ensuring the right links are on the right pages" is a problem difficultly solvable or quite a utopia. In fact, adding more levels of hierarchy in the access structure to avoid information overload may have many negative consequences.

The first one is that the path necessary to reach a given content is longer, and therefore, the number of clicks needed is higher: users may likely get lost, losing the information scent, or be unsatisfied since they cannot get the information quickly.

The second negative consequence is an increase of the number of links in the whole information architecture, and therefore, in the probability of making semiotics design errors, such as overlap in meaning or lack of clearness of the signs employed.

In this document we are trying to recall some predefined and practical search patterns and underlining their specification, which can incorporate with other dimensions of web design and helps solve critical issues. The most significant of them, called facet search takes us to many relevance advantages. Unlike a simple hierarchical scheme, faceted classification gives the users the ability to find items based on more than one dimension. Faceted search lets users refine or navigate a collection of information by using a number of discrete attributes. A facet represents a specific perspective on content that is typically clearly bounded and mutually exclusive. The power of faceted search lies in the ability of users to create their own custom navigation by combining various perspectives rather than forcing them through a specific path.

Experiences of frustration and disappointment are everyday happening when interacting with complex websites .there is no particular evidences that can connected such incidents to information architecture, search or any other aspects of this paradigm. Search alone, however, poses significant drawbacks: it does not easily provide for “serendipity”, “at a glance” sense making, “communication strength”, and “branding”. Most importantly, search can help users only when they know quite precisely what they are looking for.

In this work we advocate the adoption of rich interfaces to blend engineered Information Architectures with advanced search patterns (such as faceted search) in a seamless user experience. Our goal is to obtain effective designs both for users (achieving usability) and designers (achieving the desired impact on the user).

2 Traditional Information Architecture

One of the main goals of information architecture is to organize an informational domain into a usable taxonomy. This is, however, a difficult task: final users can classify the same domain differently from experts, differences can arise between different groups of users, and the same users can create different taxonomies for different goals (goal derived taxonomies).

According to [1] the Information Architecture is

1. The combination of organization, labeling, and navigation schemes within information system.

2. The structural design of an information space to facilitate task completion and intuitive access to content.

3. The art and science of structuring and classifying web sites and intranets to help people find and manage information.

4. An [emerging] discipline and community of practice focused on bringing principles of design and architecture to the digital landscape.

Building architects must consider the paths that occupants will follow between a building's functional areas. Information architects must do the same. When doing so, information architects must optimize a variety of different systems and data sources to satisfy their audience's data retrieval and analysis needs:

Traditionally, information architecture has been focused on the information systems that they can impact directly. The natural consequence of available technologies led architects to focuses his attention on distribution pathways within the organization. This made sense, as a distribution pathway centralizes the control and design of information entry/acquisition and information distribution. In a utopian world, such an approach allows the architect to design workflows, business processes and user interfaces to optimize for insight productivity.

A web site, in all domains, is meant to support a variety of communication goals [2], like:

- Providing practical information,
- Offering an "at a glance" understanding of what the permanent collections are about,
- Supporting a pleasant and enticing exploration,
- Allowing the user to locate a specific piece of content, as well as
- Promoting the institution's brand, some selected pieces of content (shop-window effect), etc.
- Promoting the institution's brand, some selected pieces of content (shop-window effect),

As long as the site is small, "traditional" information architecture can cope with these needs. The site gets large and information-intensive, the traditional structure starts "cracking" as layers upon of navigation are added, and disappointment becomes a common user experience.

Straight search engines have provided a reasonable solution to support just one of the above goals: Allowing user to locate a specific piece of content.

Studding current design models they are sharing a common feature: they are all based upon an information-navigation paradigm to describe the user interaction. Especially in the field of web application design [3], this legacy is due to the conceptual background underlying the origins of the World Wide Web, which derive from the Hypertext and the Data Base field: a network of links interconnects pieces of information (nodes). In this perspective, current design models consider interaction design as the activity in which the application behavior and the interface/navigation components are described.

2.1 Main elements (recall IDM) and point of strength

Users don't notice the information architecture of a site unless it isn't working. When they do notice good architectural features within a site, they instead attribute these successes to *something else*, like high-quality graphic design or a well-configured search engine. Why? When you read or hear about web site design, the language commonly used pertains to pages, graphic elements, technical features, and writing style. However, no terms adequately describe the relationships among the intangible elements that constitute a web site's architecture.

The elements of information architecture--navigation systems, labeling systems, organization systems, indexing, searching methods, metaphors--are the glue that holds together a web site and allows it to evolve smoothly. To a novice, this terminology is not very clear. These elements are extremely difficult to measure, and therefore even harder to compare. You really have to spend time using a site and get a feel for it before you can confidently talk about a site's information architecture.

The interaction of a user with an interactive application can be assimilated to a dialogue. [4] The “turns” of the user are the “interactions” (e.g. clicking over a label or pressing a button); the turns of the machine are the “response” to the user interactions (e.g. a page displayed, a file audio played).

Example: A user accessing the website of an exhibition can be assimilated to a user having a dialogue with the “curator” about the exhibition; as the user starts the application, the “machine-curator” replies with the home page, i.e. the overall synthesis of what can be said about the exhibition. The user turn (interaction) consists in selecting the wished “chapter”; the “machine-curator” turn consists into offering the content of the chapter, and so on.

The interactive dialogue model (IDM), a design model specifically tailored for multichannel applications. The model offers a set of simple primitives to describe the user-application dialogue that has to be supported to meet the requirements. IDM is particularly suitable for content-intensive or information-intensive interactive applications, whereas the major reward for the user experience is a rich set of content and messages to explore.

It is lightweight, providing a few set of primitives (and a simple graphic notation) which are easy to learn and teach. Moreover, it is suitable for brainstorming and generating ideas at early stage during design (or during the shift from requirements to design); finally, it is cost-effective (it requires little effort from designers) and modular (designers can take the part they wish, not being forced to "all or nothing").
Design can start in a "conceptual," channel-independent fashion, and then proceed into a further "logical" design oriented toward specific channels of communication.

Conceptual Design

Conceptual Design is the first step of the design process. Basically it involves the following decision: strategy about content and gross decision about info-architecture. It does not take into account the technical features of a channel, but it takes into account the context of usage.

The result of the activity is a "conceptual design document".

Below are listed the conceptual design base concepts.

Topic

Definition: A possible "subject" of a dialogue. A topic is a semantic concept, not a syntactic one.

Examples: "Michelangelo", "Leonardo da Vinci", "Galleria degli Uffizi" are examples of "topic". We can talk of them in a dialogue. "Chapter 1" is not a topic instead, since it is a syntactic concept, with no semantics; it represents a way to organize a dialogue, not a subject for the dialogue itself.

Multiple Topic (also said Kind of Topic)

Definition: A category of topics (to be kept distinct from a group of them). Some topics are **instances** of kinds of topic.

Examples: "artist", "city" are examples of multiple topics. "*Michelangelo*" and "*Leonardo da Vinci*" are examples of instances of "artist"; "*Florence"* and "*Rome"* are examples of instances of "city".

Single Topic

Definition: A topic that is not an instance of a multiple topic. The distinction is arbitrary, and it depends upon the context of the application, as explained with the examples.

Examples: "Galleria degli Uffizi" is a single topic, if it is the only museum being considered in the application. If several museums are being described, instead, than there will be a multiple topic *"museum"* and *"Galleria degli Uffizi"* would be an instance of it.

(Relevant) Semantic Relationship

Definition: A "semantic reason" to relate two multiple topics. Only "relevant" (for the sake of the application being designed) relationships must be specified, not all the possible ones.

Examples: "painted by" is a possible semantic relationship between "*painting"* and *"artist*"; "*born in"* can be a semantic relationship between "*artist"* and *"city"*, and *"active in"* and *"died in"*, can be different semantic relationships between the same multiple topics.

Inverse Semantic Relationship

Definition: A semantic relationship relating two kinds of topics, in the "opposite sense" with respect to a previous definition.

Examples: "has painted" is a possible semantic relationship between *"artist*" and "*painting"*, inverse with respect to *"painted by"*; "*birth place of"* can be a semantic relationship between *"city"* and "*artist"*, inverse with respect to "*born in"*.

Group of topics

Definition: a group of topics is a set of instances of a multiple topic.

Examples: let us assume that *"artist"*, *"painting"*, *"city"* are defined as multiple topics.

"All the artists", *"All the paintings"* and "*all the cities"* are examples of groups of topics, representing all the instances of the respective multiple topics.

"Italian painters", "Mannerists", "Impressionists" are other examples of groups of topics holding together some of the instances of "artist"; *"Countryside", "Portraits"* are examples of groups of topics for "painting"; *"Big cities"*, *"Villages"* are examples of groups of topics for "city".

"*Masters"* is another example of group of topics for "artist";"*Masterpieces", "Director's choice"* are example of groups of topics for "painting"; *"Cities of charm"* is an example of groups of topics for "city".

<u>Parametric Group of topics (also said Multiple group of topics or Kind of group of topics)</u>

Definition: a parametric group of topics represents "a set of similar group of topics". One parameter is used to "instantiate" the different groups that are similar as far as purpose and organization is concerned. Parametric groups of topic have two advantages:

- They simplify the design (1 definition instead of several ones)
- It allows delaying design details (e.g. how many countries should be consider? See the example below).

Examples: "*Painters of a Country"* is an example of parametric group of topics for "artists"; "country" is the parameter that can be instantiated in Italy, France, Germany, ..; the corresponding groups of topic would be *"Painters if Italy", "Painters of France", "Painters of Germany", ….*

"Paintings of a Subject" is an example of parametric group of topics for "painting"; the parameter can be instantiated as "*portraits", "countryside", "religious subject", "still life"* …

Logical (Channel) Design

Logical design is the second step of the design process. Basically it involves the following decisions: fragmentation of the content (taking into account several factors, including the delivery features of a specific channel), operational identification of all the possible dialogue acts, refining the info-architecture. The level of details can change, according to the specific needs of the project. The result of the activity is a "logical design document".

The following paragraphs introduce L-IDM basic concepts.

Dialogue Act

Definition: a dialogue act is a specific "concrete" turn of the machine (while a "topic" is a more abstract turn"); the dialogue act can't be partially uttered: all or nothing is delivered. It can be, however, interactively "played". The analogy is with interactive music: if user ask for a track, (s)he gets all or nothing of it; then it is up to the user to play it, either totally or partially.

Example: the user gets a web page; conceptually (and often also physically, depending upon technology) (s)he gets all of it. Then something is visualized on the screen; other parts (e.g. sub-windows) may be visualized on request; content can be interactive (e.g. zoomable images) or it can be "playable" (e.g. audio files or video files).

Content-Dialogue Act (content act)

Definition: a content-dialogue act is a specific "concrete" turn of the machine where the main purpose is to *convey content* to the user. A content-dialogue act delivers part of all the content associated to a topic (as defined in the conceptual design). Therefore, once the "channel" of delivery is defined, both single topics and multiple topics must be fragmented into content-dialogue acts (see "fragmentation" strategy).

Example: assume that for a "painting" we want to use as content a general description, a large image, a number of images about details, a bibliography, the traceable history of the painting, etc. The channel is the web; all the above content would be excessive for 1 page, therefore we decide to fragment it in several content acts.

Transition-Dialogue Act (transition act)

Definition: a transition-dialogue act is a specific "concrete" turn of the machine where the main purpose is to allow the user to change subject (transition), following a semantic relationship (defined in the conceptual design). The most typical situation is that the user wants to switch his/her interest from a topic to another, related topic, but there are several ones to choose from. Transition acts are basically lists of related topics: each possible topic must be described at some extent.

Example: the user is focused upon a content act describing a technique (for producing prints): the transition act for looking at prints created with that technique consists in a list of those prints. The user is focused upon the content act describing the curriculum of a professor: the transition act for looking at the courses thought by that teacher consists in a list of courses.

Introductory-Dialogue Act (introductory act)

Definition: an introductory-dialogue act is a specific "concrete" turn of the machine allowing the user to enter a group of topics (defined in the conceptual design). The "body" of an introductory act is a list of the topics belonging to the group. An additional optional part can be the "introduction", i.e. a piece of content used to "introduce" the group of topics.

If a multiple (parametric) group of topics is considered, there are several introductory acts (often combined in the same page): a list of the possible parameter values and the lists of the topics corresponding to teach parameters value.

Example: the conceptual design has a group of topics "masterpieces" (for the multiple topic "painting"): the corresponding introductory act must contains a list of paintings (those selected as masterpieces). Optionally there can be an introduction, possibly explaining the concept of "masterpiece", why those paintings have been selected, who selected them, etc.

The conceptual design has a parametric group of topics "paintings by subject", holding together all those paintings related to a given subject. There must be an initial introductory act listing all the possible subjects (and an optional introduction); to each subject will correspond an introductory act listing the paintings (of that subject) and with a possible introduction.

Fragmentation (of topics)

Definition: If a topic (from conceptual) design is small (in terms of content size) it can be "said" with one content act; if it is large, however, it must be fragmented into several content acts. This fragmentation can be decided with different criteria: by size (in order to keep each fragment simple), by content (by separating different types of content), by user profiles (keeping in different fragments content suitable for different users), by media (separating "normal" media, such as text or low quality images from more demanding media such as high quality media, video, audio, etc.), ….

Important: the different fragments may duplicate information (and very often they do it).

From the above it derives that the fragmentation must keep into account the specific features of the channel, and also the context of usage.

Example: consider that we have a large set of information for each painting; here are some of the possible strategies for fragmentation:

- Fragmenting by content type: content acts (one or several ones) with technical information, content acts (one or several ones) with historical information, content acts (one or several ones) with art comments, content acts (one or several ones) with contextual information, etc.
- Fragmenting by user profiles: content acts (one or several ones) for the "naïf" users, content acts for the "expert" user, content acts for the "teacher", content acts for the "school child, etc.

- Fragmenting by "depth": content acts (one or several ones) with introductory material, content acts (one or several ones) with advanced material, content acts (one or several ones) with reference material, etc.
- Fragmenting by media: content acts (one or several ones) with texts and low quality images, content acts (one or several ones) with high quality images, content acts (one or several ones) with virtual tours in 3D graphics, .etc.

Note: fragmentation has an impact upon structural strategy (see below) since if a topic gets highly fragmented, each fragment is easily manageable (advantage) but a more complex navigation is required to get all the fragments; if a topic is no fragmented at all or lowly fragmented, little navigation (or no navigation at all) is required (advantage) but each fragment can be heavy (on the interface and on the user).

Structural strategy

Definition: the structural strategy defines the navigation (interaction) that is used to interconnect the different content acts corresponding to the same topic.

If a topic (from conceptual design) is not fragmented, no structural strategy is involved. If, instead, several content acts are devised, then there must be a way, for the user, to navigate across them. The structural strategy must define the following:

- The *initial (default) act*: i.e. the content act first offered to the user
- The *interaction (navigation)* for reaching the other content acts of the same topic

Typical navigation patterns are *all-to-all* (i.e. from any content act is possible to reach any other content act), or *tree* (i.e. from the initial act, all the others are reachable directly) or *guided tour* (i.e. form the initial act the second act can be reached, and then the third, etc.).

Structural strategy is defined very roughly at logical design time, since the details of the navigation are decide at page design time.

Transition strategy

Definition: the transition strategy defines the navigation (interaction) that is used to interconnect the different topics corresponding to "source" and "target" of a semantic relationship, the source being the topic from where the relationship originates and the target being the topic where it lands.

If the target identifies a unique instance, there is no transition strategy. If the target identifies several possible instances, then a list of possible instances is needed.

The transition strategy must define the following:

- The *interaction (navigation) pattern* used *for* reaching the instances of the target topic.

Typical navigation patterns are *index* (i.e. a list of possible instances is offered, and the user must select one of them), *guided-tour* (i.e. the first instance is offered, and then next one, etc.) *or mixed* (i.e. index combined with guided tour).

In both cases orientation information (see page design) is strongly needed.

Introductory strategy

Definition: the transition strategy defines the navigation (interaction) that is used to interconnect the different topics corresponding to "source" and "target" of a semantic relationship, the source being the topic from where the relationship originates and the target being the topic where it lands.

If the target identifies a unique instance, there is no transition strategy. If the target identifies several possible instances, then a list of possible instances is needed.

The transition strategy must define the following:

- The *interaction (navigation) pattern* used for reaching the instances of the target topic.

Typical navigation patterns are *index* (i.e. a list of possible instances is offered, and the user must select one of them), *guided-tour* (i.e. the first instance is offered, and then next one, etc.) or mixed (i.e. index combined with guided tour).

In both cases orientation information (see page design) is strongly needed.

Organizing a list

Definition: there situations when list are needed, i.e. transition acts and introductory acts. In both cases list must be carefully designed, in order to optimize effectiveness and usability.

When a list is needed, two main decisions must be taken:

- Which information to put for each line, i.e. which information should be used in order to identify each instance referenced by the list.
- How to organize the sequence of lines, i.e. in which order to put them.

The need for being as much as possible "informative", would dictate to use several information items for each line. At the same time the need to be simple and to "pack" several lines in a small space would dictate to use a few items for each line.

Navigational strategy

Definition: the navigational strategy, i.e. the way to navigate across dialogue acts, is related to the perception of the user ability and interest.

Generally speaking a user with little previous knowledge of the content or using an application for the first time would prefer simple (possibly inefficient) strategies (e.g. guided tours). An expert user would rate efficiency over simplicity, and therefore prefer indexes.

Long lists need effective strategies such as "filtering" or "reordering" (see advanced features) and are not suitable for guided tours.

Page Design

The page (interface) design completes the design process. Basically it takes the following decisions: transformation of dialogue acts into pages (or equivalent notion for some channel), completing the details about navigation and links, identifying the proper elements of the interface (e.g. name of objects, name of links, names of lists, etc.).

The result of the activity is either a mock-up or a prototype or the real application, accompanied by explanatory documentation.

2.2 Points of weakness

When confronted with a new and complex information system, users build mental models. They use these models to assess relations among topics and to guess where to find things they haven't seen before. The success of the organization of your web site will be determined largely by how well your site's information architecture matches your users' expectations. A logical, consistently named site organization allows users to make successful predictions about where to find things. Consistent methods of organizing and displaying information permit users to extend their knowledge from familiar pages to unfamiliar ones. If you mislead users with a structure that is neither logical nor predictable, or constantly uses different or ambiguous terms to describe site features, users will be frustrated by the difficulties of getting around and understanding what you have to offer.

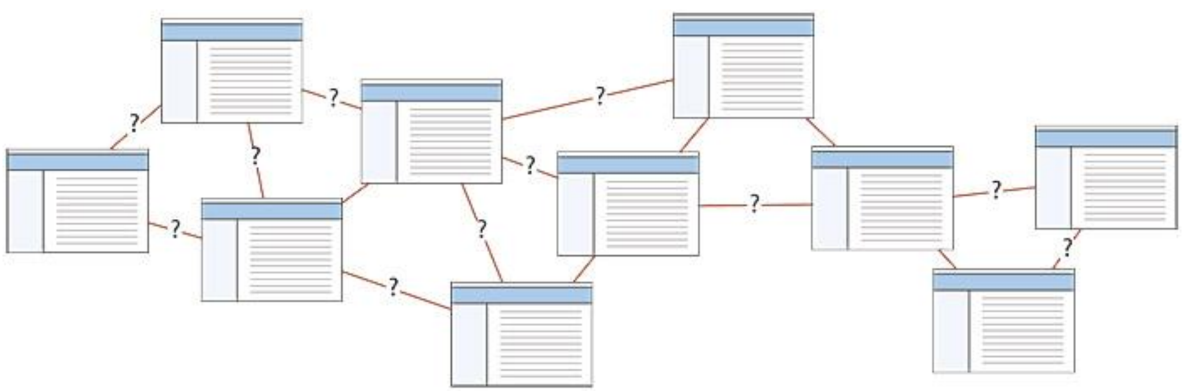


Figure 2-1 A confusing web of links. Users try to imagine the site structure, and successful information architecture will help the user build a firm and predictable mental model of the site.

2.2.1 Problem with Hierarchical structures

The goal is to build a hierarchy of menus and content pages that feels natural to users and doesn't mislead them or interfere with their use of the site but usually the Web sites with too shallow an information hierarchy depend on massive menu pages that can degenerate into a confusing laundry list of unrelated information. Menu schemes can also be too deep, burying information beneath too many layers of menus. Having to navigate through layers of nested menus before reaching real content is frustrating. Here the problem is that navigation is fixed but the hierarchy gets too deep and unmanageable.

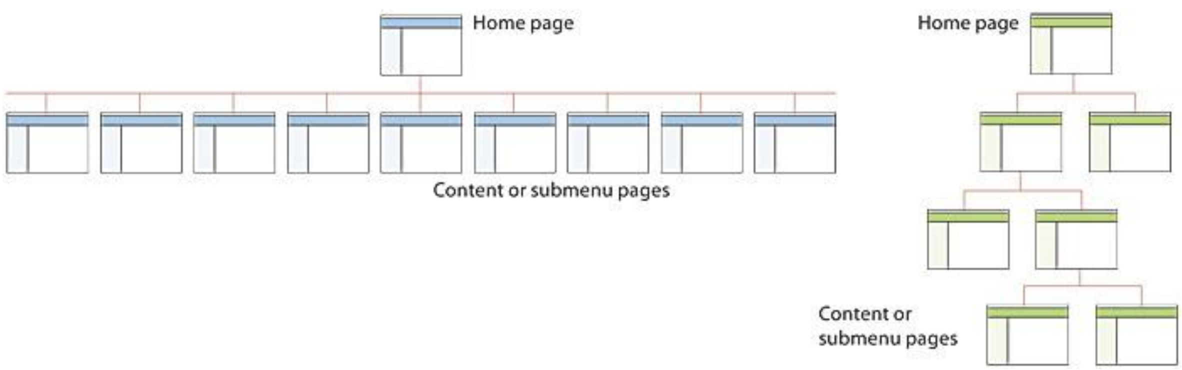


Figure 2-2 Too shallow a structure (left) forces menus to become too long. Too deep a structure (right) and users get frustrated as they dig down through many layers of menus.

In this situation there is no balance of menus and content pages and usually menu scheme hast outlived and has week areas. Because of this kind of complex structure we have deeper menu hierarchies and so users be forced into page after page of menus. So quick access to information is not possible and the organization of the site isn't reflecting clearly. this hierarchy of content subdivisions should not become a navigational straitjacket for the user who wants to jump from one area of the site to another. So in general the problems with the hierarchy are:

- Inflexible
- Force the user to start with a particular category
- What if I don't know the animal's diet, but the interface makes me start with that category?

- Wasteful
- Have to repeat combinations of categories
- Makes for extra clicking and extra coding
- Difficult to modify
- To add a new category type, must duplicate it everywhere or change things everywhere

2.2.1.1 Problems with Group of topics

Here if we arranged theses contents into a single hierarchy so we can't combine freely multiple classifications therefore we keep them separated and we can't combine multiple criteria (e.g. by subject and by year). From other perspective if we perform to combine the contents we have to "nest" them which make other navigational problems (e.g. first choose the subject, then the year).

2.2.1.2 Problem with Isolated topics

In many large website usually designers face with a great amount of non-homogeneous topics. They are single topics, and therefore cannot be clearly grouped together, so they are put here and there, very arbitrarily. The problem is to find the appropriate location for them in the hierarchy.

In Louver museum official website as there are many contents which are concerns single topics. They are collecting spars material about events like seminars or conferences which are given to the louver, but as there is no specific part for these similar stuffs so they put them in auditorium section. However users maybe couldn't find these similar contents as he doesn't expect these are inside auditorium section.

The same problems are for Books, DVDs & CD-ROMs which are inside museum section. User normally expects information like history or about inside this substance (museum) therefore they maybe couldn't find easily information about books in this place.

2.2.1.3 Problems with Contextual navigation (transition acts in IDM)

The contextual navigation could ameliorate the findability of those items whose classification has been proven to be problematic, and to increase the information scent of the whole domain. Here the utility is in the design of the navigation of the information space, to integrate the hierarchical navigation with a contextual navigation. This kind of navigation corresponds to in-line navigation and according to (Rosenfeld and Morville) contextual navigation emphasizes the creation of links specific to a particular page, document or object. They cite the "See Also" links to related products in e-commerce sites. A famous, and very useful, example of contextual navigation is used by Amazon: "Customers Who Bought This Item Also Bought".

A smart use of contextual, supplementary navigation could be a solution to improve the findability of the items with low category validity. Using a "See Also" list of links, the designer can show the ambiguous items in all the category pages where the user could search for them.

There are several problems we can mention dealing with contextual navigation

a) A contextual navigation has been proposed that should improve the findability of the difficult elements, without violating the hierarchical structure of the system. But having many relations caused same problems discussed above; such as need to hierarchies and need to manage heterogeneous instances. Therefore we should take into account few numbers of relations.
b) Here we don't have the ability to navigate at level of group of topics: we cannot say I have these artworks, give me the list of artists that painted them.

- **Items to Related Items**
    - Problem Definition

There are various situations that the readers of a Web resource need to Improve information flow throughout by providing accurate and consistent information for better communication and collaboration."Findability" is one of the major focuses for the users. All content entered into the site is tagged with metadata, which is then used to display the content in a variety of ways.

For some information tasks, it is sufficient and perhaps even natural to deal with just one such unit of web information at a time. For example, to get a basic understanding about the Politecnico students who are admitted at PhD program for current year, one can simply read the "admission list" page. However, if one wants to quantitatively compare the international students of various countries previously admitted, then he must browse many WebPages, extract out the quantities for comparison, tabulate them in a spreadsheet, and then construct visualization such as a bar chart to show the comparison.

Figure 2-3 Related items

In traditional approaches, only the relationships among concepts explicitly described in the conceptual schema are available to the user for browsing and retrieval. Therefore, all possible relationships must be anticipated and described: a very difficult if not impossible task.

That is, when dealing with data in the aggregate, in order to gain a big picture from bits and pieces extracted out from several sources, the current Web and web browser fall short due to their "one web page at a time" browsing paradigm, leaving the user to flip between several web pages and carry out tedious data tabulation manually.

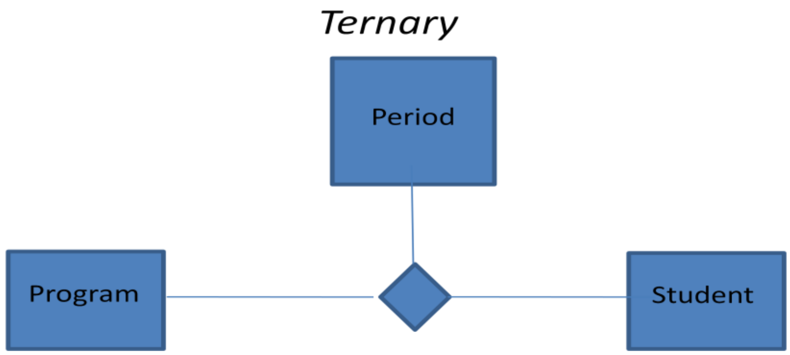


Figure 2-4 Ternary relationship

Here, you are trying to discover relationships or trends between contents and each content has too much related content to be displayed .Also through fixed navigational structures, it is very difficult to navigate and find related spread items.

The concept of relation also in IDM is in many cases too simplistic, as it doesn't model N-nary relationships, in many cases, relationships should be additionally characterized, e.g. The person - role - project is a ternary relationship having just two sides (person and project) may be too poor, we can use the third size (role) as a filter.

Dealing with one item usually you perform most of the related connections which all together can match your search. So user can easily go through any available filed and any available values which are presented in whole data .navigating should be more interactive to give these opportunities to a user.

- Navigation as a Means to Inform Relationship

As mentioned above since this is our desired relationship which we need to use and show site users that these Items, pages and concepts are related. If we do not have a strategy or navigation mechanism that include a link or way to each of the related items, we would have a more difficult time helping the user find all of the pages related to each member of the relationship.

From an SEO perspective, it's essentially achieving the same goal. It's a way to show the search engines additional related information.

Sure, we could just link from the home page to every page on the site, and from every page back to the home page. If we do, we're leaving it up to the search engines to rely solely on other indicators (keyword relationships, for example) to determine algorithmically which pages are related.

By using the controlled linking in this example, we provide a very strong additional indicator, reducing the potential for keyword based relationship determination to be off the mark.

- Adding Wight to the relationships

A relation in everyday life shows an association of objects of a set with objects of other sets (or the same set) such as Professor X teaches Multimedia Course. The essence of relation is these associations. A collection of these individual associations is a relation, such as the Teaching relation between peoples and Course. To represent these individual associations, a set of "related" objects, such as Professor X and a Multimedia course, can be used. However, simple sets such as {Professor X, Multimedia} are not sufficient here. In some cases we need to add a weight to the relationship.

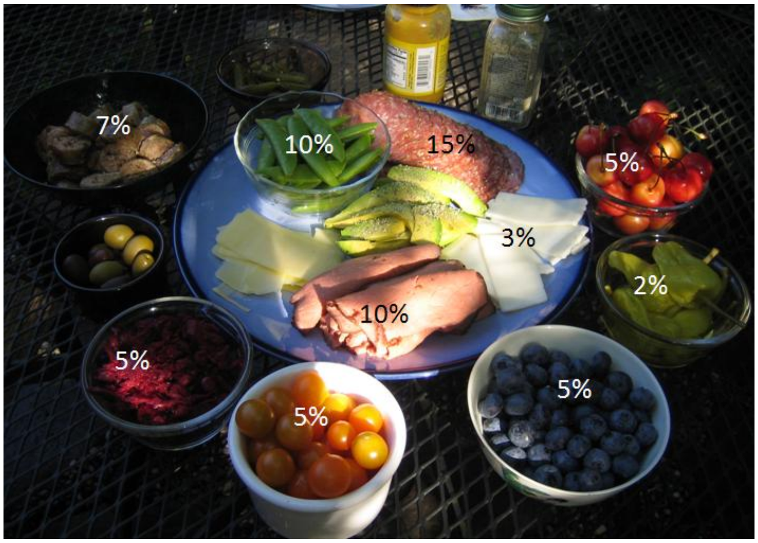


Figure 2-5 Adding weight to relationship

Navigating through websites there are several situations which user need to get information about the value which can play a third hidden dimension role. Sometimes user need to know the attributes of involvement of first item in relation with the second item. These weights will make the whole cloud of “Findability” of the user to be allowed to located structured information involving several items and make her to have a more reliable idea in specific concept.

The lack of this issue also will affect the concept of "Discovery or serendipity" of the users. He couldn't stumbles into something that he was not aware before. Concerning this aspect, we will have a generic engagement and he will browse around at pleasure.

Figure 2-6 Binary relationship

- Transitivity of relationships:

Another aspect of related to relation concept is Transitivity. A relation is transitive if for all values a, b, c: a R b and b R c implies a R c

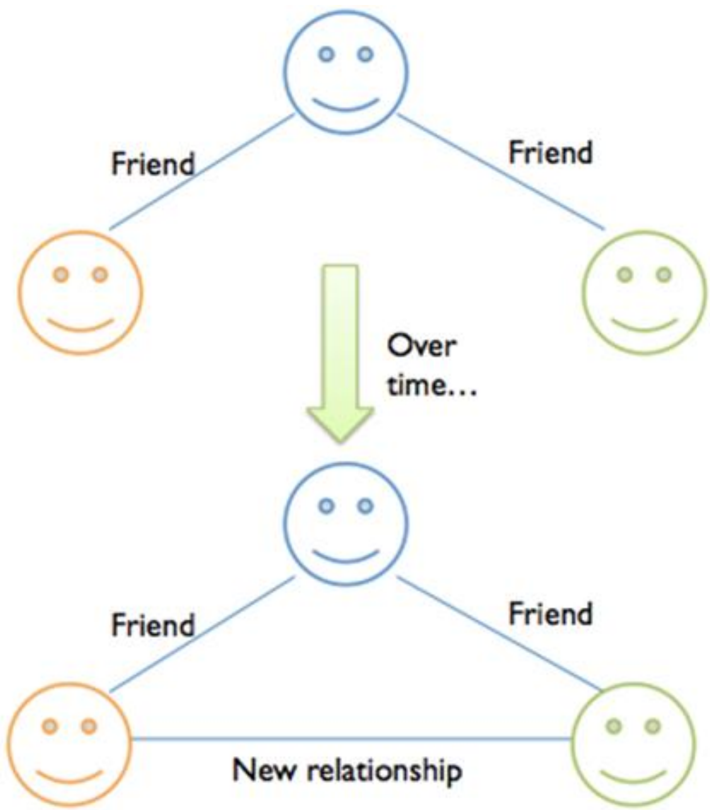


Figure 2-7 Transitivity of relationships

In social networks, transitivity means "the friends of my friends are my friends". The transposition of this property in behavioral networks implies "the users looking like those who look like me look like me". While browsing the website If a is connected to B and b is connected to C, maybe I should allow to move directly from A to C.As an example Msc of Computer science is held in Como and Professor X is teaching in Como campus so Professor X is one of the teachers of Msc of computer science.

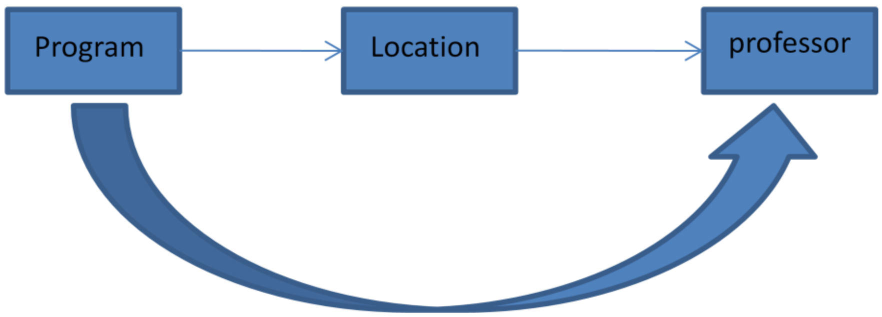


Figure 2-8 Relationship

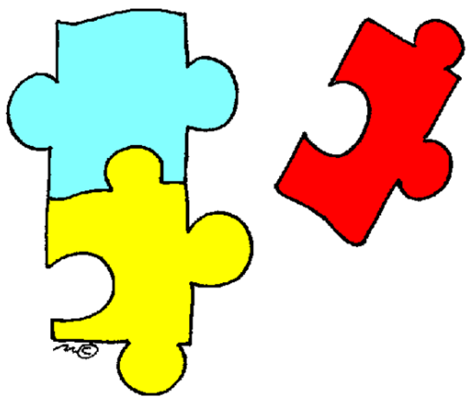
As explain above there should be perform navigating through different related items with the aiming of providing several ways of gain access to the related contents. For large websites, however, the overall hierarchy resulting from the information architecture is not completely satisfactory: the user cannot easily locate what is looking for, and even interesting information (with a great possibility of serendipitous access) is buried under the levels and levels of navigation necessary to overcome problems of information overload.

c) You cannot see the relation at multiple levels of granularity. e.g. sometimes you may be interested in who joined a certain project regardless of the role, sometimes you are interested in who joined a certain project with a certain role and in certain period.

2.3 Inference of Metadata

Another aspect which should be taking into account is inference of metadata. Users need to have permits to interactively change the sorting and grouping of elements, according to the metadata assigned to contents. Metadata is structured data which describes the structural characteristics of a resource and it usually refers as "data about data".

By clicking on links the user need to select a combination of metadata values belonging to several classifications.



I can use the English programs to access the computer science and also we can choose the thematic area as engineering to get the computer science as a result.

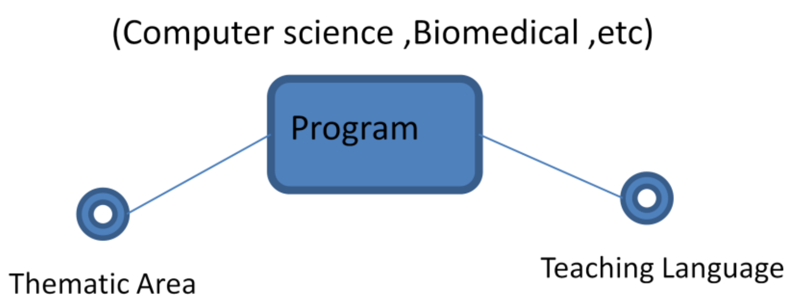


Figure 2-9 Inference of metadata

In some cases we need to use facets of the related topic Y to browse topics of type X, for example we use nationality of artists to browse artworks; we need to use the genre of tracks to browse albums ; moreover information about artworks to browse museums. So we need the high degree of connectivity to response to these particular requirements.

In general the problem is the possibility to access the content from the list related one, you need to go through each item to visit related contents. We can notice that in these cases access structures are arranged in a single hierarchy. We can observe that traditional information architecture forces the web designer to make one criterion of classification. Prevail on the others by placing it at the top level of the site structure, as starting point for navigation. User need to browse contents according to several classifications at the same time but the traditional hierarchical information architecture forces the user to choose a single classification from the ones provided.

3 Search Patterns

According to [5] Search is among the most disruptive innovations of our time. It influences what we buy and where we go. It shapes how we learn and what we believe. This provocative and inspiring book explores design patterns that apply across the categories of web, e-commerce, enterprise, desktop, mobile, social, and real time search and discovery.

Explicit navigation serves as a map for way finding and understanding, while the search box offers a shortcut that employs semantics to bypass structure. Often, these modes need be distinct only in the opening. In search, users make the first move by typing words to declare intent. But, the SERP (search engine results page) is a browsable interface with visible context. Alternatively, you may begin with browse, but then query the category that you're inside using scoped search. A well designed system lets people flow between modes and offers immediate feedback, because in the endgame, it's all about interaction.

Search is a wicked problem for two reasons [6]. First, it's radically multidisciplinary, requiring real collaboration between design, engineering, and marketing. Most organizations simply can't get these folks to work together. Second it's a project and a process, requiring a major initial investment and the commitment to continuous improvement. Few organizations are good at both.

3.1 Main search patterns

Having considered the characteristics and motivations of systems for representing information, we now turn to the problem of making that information accessible.A concept that originated in software engineering; a design pattern is a general, reusable solution to a commonly occurring design problem.

3.1.1 Textual classical search (The Keyword Search)

A popular solution has been to input a keyword for searching and build a traditional relational database query. This was straightforward to implement if: (i) the data already resided in a relational system, (ii) the search was simple, and (iii) the author of the system was knowledgeable about the database.

Unfortunately, as systems like this deployed, it became increasingly more difficult to build the required commands to return the correct result. An additional problem arose for users conducting keyword searches -- they were unable to receive feedback indicating which aspect of their query caused the search to leave them empty handed. Essentially, this method of finding information tends to break down quickly when both the amount and complexity of data increase, and can be thought of as a guess-and-pray (G.A.P.) search. With a keyword search there is simply no guarantee of finding information immediately or knowing with certainty that it doesn't exist.

Case in point: an individual shopping online at an office supply site is looking for a shipping option and is given the opportunity to conduct a keyword search. He or she enters "FED-EX" for Federal Express. The content creator of the office supplies catalog might have entered FED Ex, Federal Express, or fedex when creating the site. Therefore, unless the individual guessed the correct keyword, the search would fail [8].

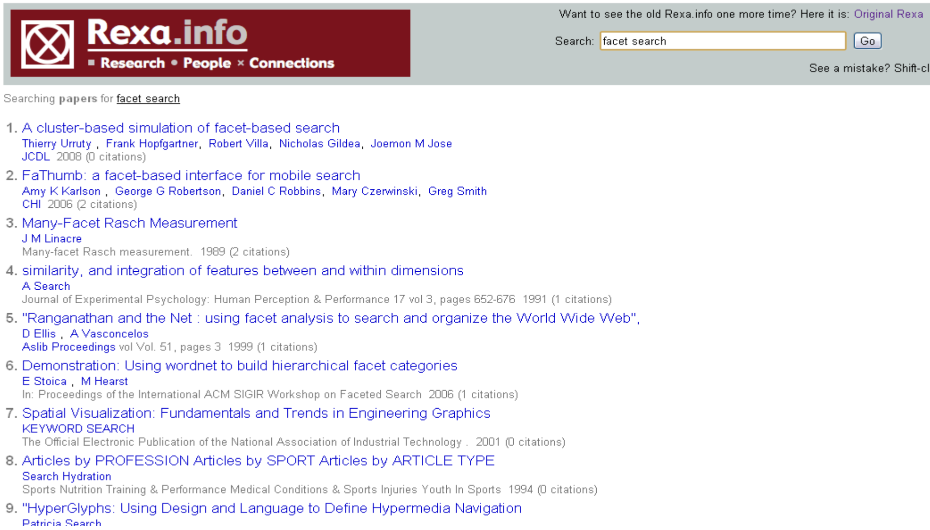


Figure 3-1 Keyword search example

There are two main kinds of queries: structured queries and text queries. Structured queries work on structured data (e.g. relational DB tables and fields, XML, or the semantic web...). They are very precise but they also require learning a specific syntax (SQL, XPath, SPARQL...).

Text queries just work on text. The search engine try to guess if a certain piece of content (a document, a page, an entry in the database, etc.) is somehow related to the query terms you entered. You may have undemanding text-based search that simply checks if the terms you enter are inside the documents you are searching for. Or you may have very sophisticated forms of search that involve some form of data mining, natural language processing, etc.In any case, text queries are uncomplicated search patterns to work with but are more likely to produce unsatisfactory results.

Free-text queries are, without a doubt, easier for users to create than formal Boolean expressions. But this liberation comes at a cost: the query no longer represents a well-defined filter on the document collection. Instead, the query becomes a target, and the binary notion of a document matching the query relaxes into a continuous similarity measure [9]. An information retrieval system built on such an approach no longer filters documents but rather ranks them according to the degree to which they match the query.

3.1.2 Instant suggestions

As users type into a text entry box, suggestions appear automatically. It's a pattern that first emerged within the Help functions of desktop applications. This solution solves several very common problems. First, typing takes time. Second, users can't spell well. Third, we're often at a loss for words. We simply don't know or can't remember the right terms.

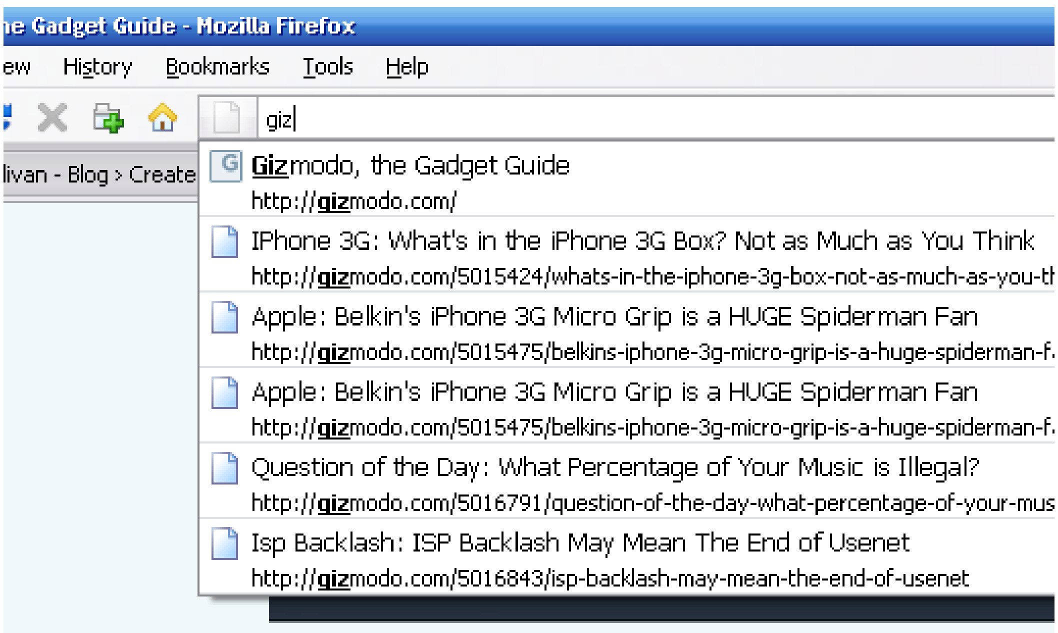


Figure 3-2 Firefox instant suggestion

Since these problems are portable, it was only a matter of time before the pattern spread. Auto complete is now a familiar fixture across desktop, web, and mobile platforms.

A major prerequisite of autocomplete is a source of data for suggestions. Classic desktop applications rely on an alphabetical index of help topics. Google draws from a user's personal search history and from the collective search behaviour of many users. Firefox taps browsing history and bookmarks.

The textbox for entering queries could have an interesting added functionality: an instant suggestion of keywords to search, determined on the basis of what the user is typing. An example of this widget is the one of Google Suggest.

(http://www.google.com/webhp?hl=en&complete=1).

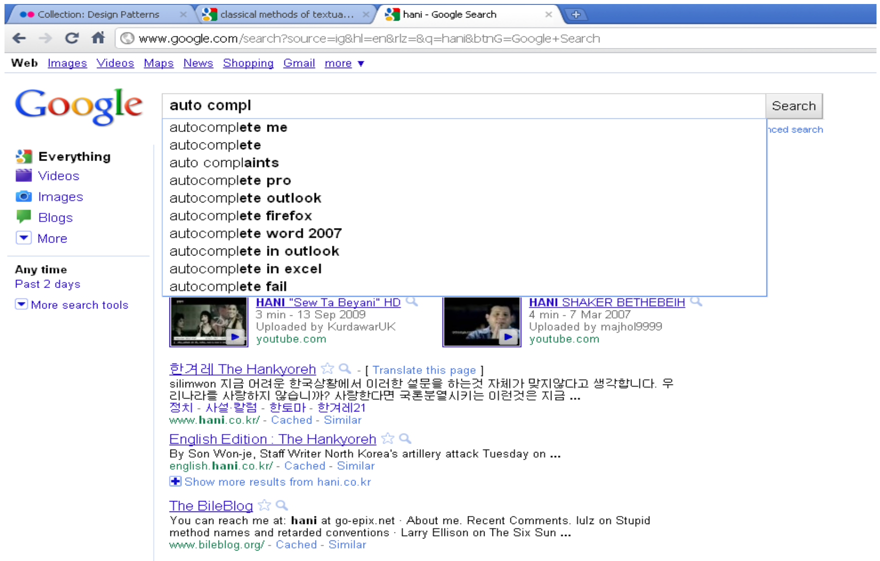


Figure 3-3 Google suggestion search box

An instant suggestion box brings two important advantages:

It prevents user misspellings, since user can select the appropriate query from the suggested ones without type it entirely. It reassures users about the appropriateness of their query and consequently the possibility of obtaining relevant results. In fact, if others (supposed many other users) have typed the same keywords, those keywords would more likely bring to the desired information.

As for the list for most frequent searches, the suggestion box could rely on the search engine statistical records. For each suggested query, it could report also the number of obtainable results: that could be a significant index for the user to understand if his/her search will produce useful results.

Like for Google suggest, it could be implemented using AJAX, which uses JavaScript to manage the requests to the server (in this case, for obtaining suggestions) and change single elements of the page. The application should be fast enough to react properly to user's actions, which consist in:

- Typing and deleting already typed characters
- Selecting one item from the list of suggested queries.

A smart solution to speed up the behaviour of the search box could be caching already acquired suggestions: in this way, if the user deletes a character, it is not necessary to request them again from the server.

3.1.3 Faceted search

Faceted or "guided navigation" leverages metadata fields and values to provide users with visible options for narrowing or refining their query. This integrated searching and browsing experience lets users begin with a simple keyword search, but then offers useful next steps and insights into the content and its organization. Guided navigation has become almost ubiquitous in e-commerce given the availability of structured metadata and the clear business value of improving findability. This pattern is applicable in most contexts, assuming a substantial volume of valuable content. This pattern needs human effort for defining metadata fields and values and tagging can be distributed and/or semi-automated. A problem with assigning documents to single categories within a hierarchy is that most documents discuss several different topics simultaneously. Text consists of abstract discussions of ideas and their inter-relationships. It is a rare document that is only about trucks; instead, a document might discuss recreational vehicles, or the manufacturing of recreational vehicles, or for that matter the trends in manufacturing of American recreational vehicles in Mexico before vs. after the NAFTA agreement. The tendency in building taxonomic hierarchies is to create ever-more-specific categories to handle cases like these. A better solution is to describe documents by a set of categories, as well as attributes (such as source, date, genre, and author), and provide good interfaces for manipulating these labels.

This use of what is known as faceted metadata provides a usable solution to the problems with navigation of strict hierarchies. The main idea is to build a set of category hierarchies each of which corresponds to a different facet (dimension or feature type) that is relevant to the collection to be navigated. Each facet has a hierarchy of terms associated with it. After the facet hierarchies are designed, each item in the collection can be assigned any number of labels from the facet hierarchies. The resulting interface is known as faceted navigation, or sometimes as guided navigation. An alternative design, usually known as parametric search, requires the user to select a number of attributes from drop-down menus all at once, thus often leading to empty results sets. (For instance, at a Web site selling shoes, selecting women's shoes of size 4.5, width M, color black, and price to yields no results and provides no suggestions as to which attributes to remove.) For this reason, parametric search has largely fallen out of favor and been replaced by faceted navigation.

Faceted classification and faceted navigation are now widely used in web site search and navigation. With this kind of search we can reach flexible navigation, seamless integration of browsing with directed (keyword) search, fluid alternation between refining and expanding, avoidance of empty results sets, and at all times allowing the user to retain a feeling of control and understanding.[7]

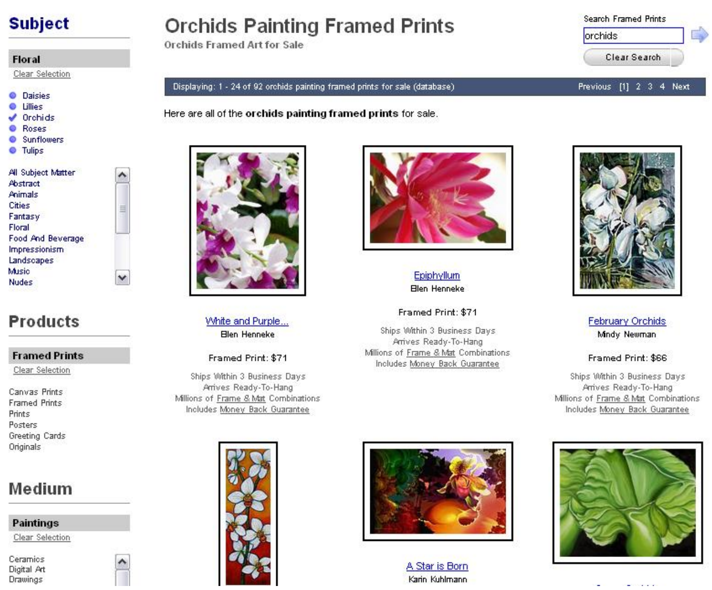


Figure 3-4 Flamenco interface for browsing and search using hierarchical faceted metadata. Item view showing the leaf-level metadata, placed in the context of their facets

Faceted search addresses weaknesses of conventional search approaches and has emerged as a foundation for interactive information retrieval. User studies demonstrate that faceted search provides more effective information-seeking support to users than best-first search. Indeed, faceted search has become increasingly prevalent in online information access systems, particularly for e-commerce and site search. We can use facet navigation in online systems, as an alternative to the hierarchical focus of web site structure, and in response to the failure of subject searching in online catalogues.

According to Hearst et al., 2002 describe faceted metadata as being composed of "orthogonal" sets of categories, meaning each category describes a different, usually independent aspect of the information items. For example, in the domain of fine arts images, possible facets might be Media (etching, woodblock, ceramic, etc.), Locations (Asia, North America, Europe, etc.), Animals & Plants, Earth & Sky (mountains, rivers, clouds, etc), as well as information about the art, including artist names, time of creation, etc. A facet may be flat ("by Pablo Picasso") or hierarchical ("located in Vienna > Austria > Europe"). A facet may be single-valued or multi-valued. That is, the data may allow at most one value to be assigned to an item ("measures 36 cm tall") or it may allow multiple values to be assigned to an item ("uses oil paint, ink, and watercolor"). Portions of the hierarchies within a facet are sometimes referred to as the facet's subcategory or sub hierarchy.

There are some deficiencies with the faceted paradigm. If the facets do not reflect a user's mental model of the space, or if items are not assigned facet labels appropriately, the interface will suffer some of the same problems as directory structures. The facets should not be too wide nor too deep (with exceptions for long lists such as author names that cannot be organized meaningfully) and the interface must be designed very carefully to avoid clutter, dead ends, and confusion.

Some points cope with facet search are:

- The power of faceted search can overwhelm and confuse users if it is implemented with a poor design.
- Choosing the correct layout of facets and results in an interface can be a trade-off between the work required for users to see results and the likelihood they notice the facets.
- For ambiguous queries, users may benefit from a facet-driven clarification dialog.
- Strategies to avoid information overload by filtering facets and facet values also offer ways to rank and organize them.

- Users generally expect that initiating a new free-text search will clear current query filters, but some applications provide the option to search within current results.
- A number of decisions about search behaviour involve a precision/recall trade-off: consider computing results to favour recall but computing the utility of facets and facet values based on a narrower search query that favours precision.
- An interface allowing multiple selections within a single facet should not only be self consistent but adhere to familiar conventions.
- Consider the use of design patterns to take a holistic approach and learn from the collective wisdom of practitioners.

3.2 General disadvantages of search

Searching for information in large rather unstructured real world data sets is a difficult task, because the user expects immediate responses as well as high-quality search results. Today, existing search engines, like Google, apply a keyword-based search, which is handled by indexed-based lookup and subsequent ranking algorithms. This kind of search is able to deliver many search results in a short time, but fails to guarantee that only relevant data is presented. The main reason for the low search precision is the lack of understanding of the system for the original user intention of the search.

According to [12] Search engines – both external or within the website – are often the only way or "the last attempt" for the user, to find what is looking for. Nevertheless, a total reliance on traditional textual search (e.g. on the style of Google) is far for being an optimal solution, for a number of reasons: the user may have a generic need, difficult to translate in a specific search query; moreover the overall communicative "message" promoted by the website might get lost. In other words, the balance between push (contents that are offered by the website without explicit demand) and pull (contents accessible only on demand) would be too much moved towards pull.

3.2.1 Lack of introductory content

When relying on search context there is a lack of introductive information which is easily implement able by traditional hierarchical navigation structures. The user is presented with many choices, and has not clear guidance on how to start.From this point of view user encounter with a lot of data and more powerful options to filter them and gain the information. Nevertheless there is no real “guidance”. In facet search users have to practice with the parameters used for filtering. And no one explains them and no hint ensures that the query they are entering suits their "inner" research question. Moreover, usually the system requires its users some knowledge about the particular subject. Thus, concerning this point of view, faceted search lacks the richness of an organic discourse on the topic offered for instance by the IDM information architecture framework. In some cases like in artistrising website [16] which is a collection of fine art and photography we can see this problem implicitly.To underline this problem of faceted navigation with respect to traditional information architecture, we can have one realistic scenario of usage for the artistrising website.

For an assignment a student in Photography needs information about digital photos, artists and styles in Canada for understanding and analyzing what are the most famous collection and who are the famous digital photographers in Canada with respect to digital photography in this website which is one of the source of items and arts of this kind.

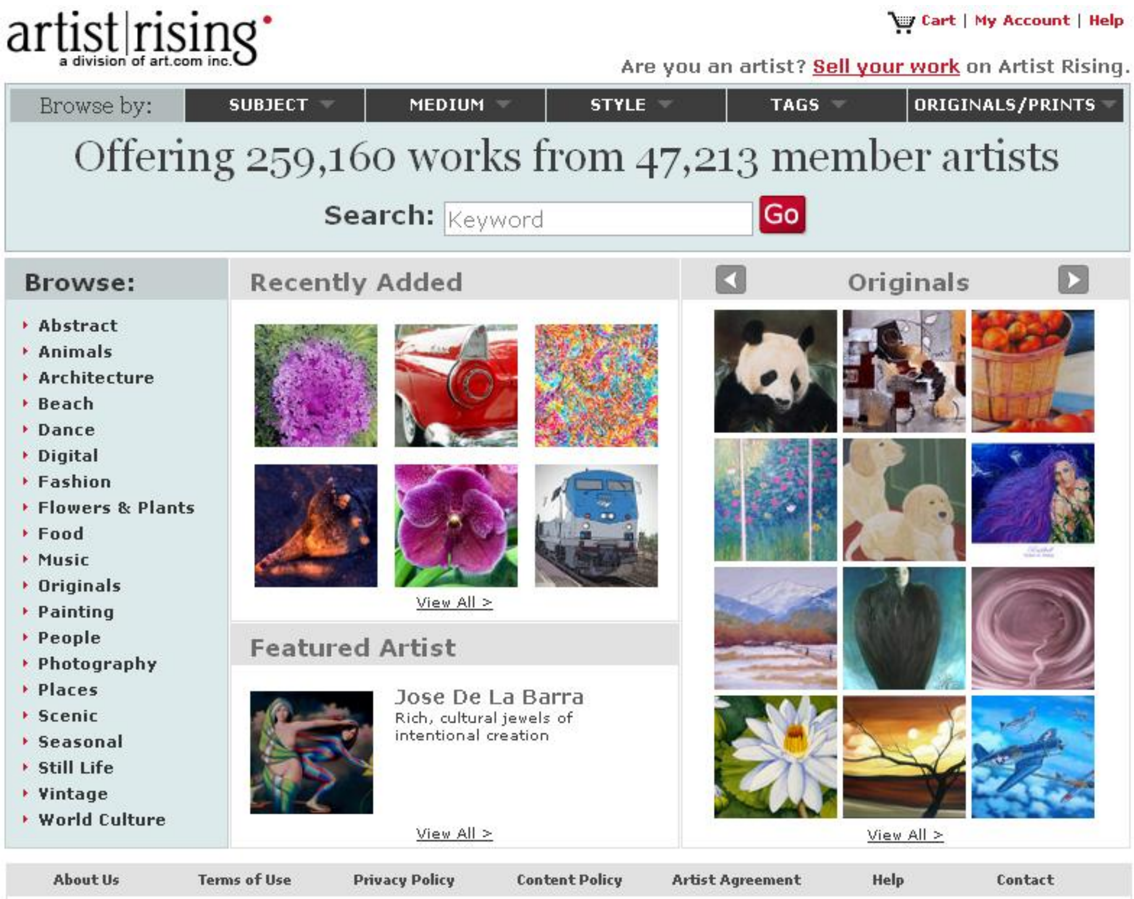


Figure 3-5 Artisrising website example for introductory content

However, we notice that in a paradigm totally relying on faceted search, there is no information on digital photography in general: what they are, who are their artists. Moreover, in some cases, like the one of this scenario (in which the user is interest only in digital photographers, artists and anything related in Canada and not in photo galleries and artistic sites) the user is provided with a superabundance of choice options deriving from facets she do not care of. This may generate a form of information overload that is different from the one of single hierarchies, but might be equally dangerous for the user experience.

3.2.2 Lack of structured navigation

Another crucial aspect should be taken into account dealing with search aspects and search patterns are the general structure and navigation. Navigating through different contents dealing with facets and several filtering options should not take us to a motion that a site loses all sense of structure and organisation. What happens if the user selects a certain result to view it in detail? How user can navigate through search results and going back to them ensuring the orientation, i.e. the user awareness or the current context and status of navigation? If such problems have been widely elicited for traditional web architecture, they are not clear enough for faceted navigation.

A good real example which illustrates this problematic issue is "freebase"[17] website. There is a part "filter this collection" which is not any assistance for users to find and understand how they can work with and how to navigate through the search results after using this option.

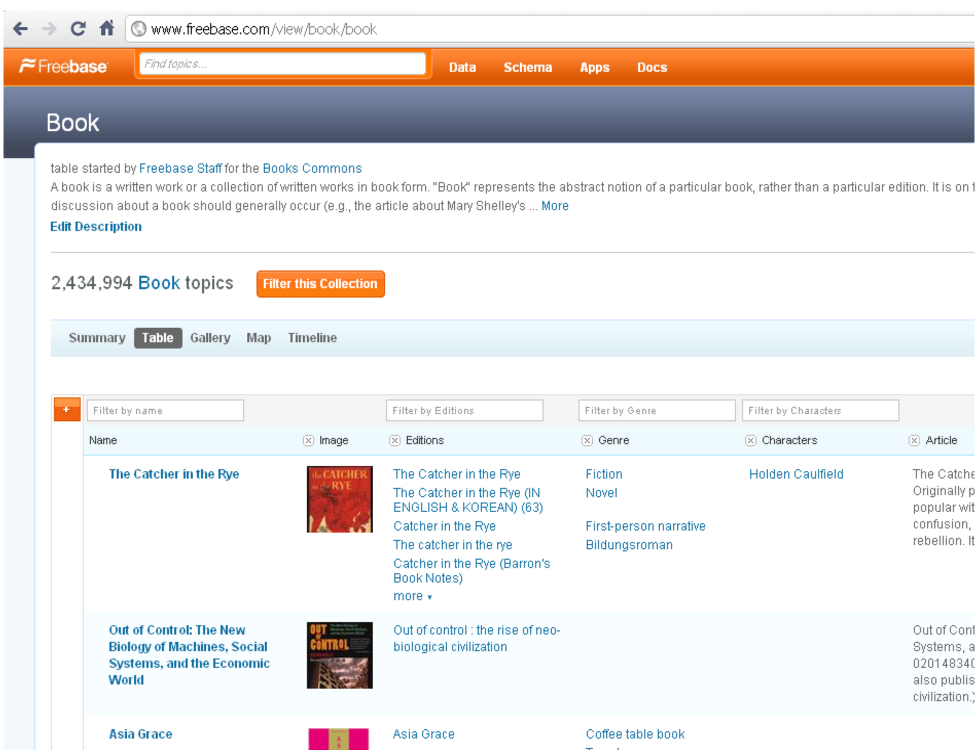


Figure 3-6 Freebase website lack of structured navigation

We can emphasize that search has been conceived, and is still conceived, for advanced knowledge needs of users who (1) are usually (somehow) experts of the field (2) are motivated to learn how the system works, to take advantage of the most powerful features.

As an example a student who is attempting to write her thesis browsing on some scientific database e.g. Web of Science or Google Scholar to look for interesting papers in her thesis area is expert of her area, so she knows, or imagine knowing, the keywords she should use. She is motivated to learn how the web application works, because she *must* find some information to write her thesis. This is definitely not the case of a common website user. We cannot expect that, a user is disposed to learn how a Tourist Office website works. Therefore it should be intuitive and provide some guidance; otherwise user simply quits and goes to another website.

That is the main difference between search-based professional digital archives and IA-powered communication-intensive websites. The challenge is to take some search functionalities and bring them in the IA world (also simplifying some patterns with respect to their "advanced" version), to support new different user experiences.

3.3 Important issues of search

3.3.1 Relevance

The core concern of information retrieval is to help users retrieve documents that are relevant to their information needs. The notion of relevance, however, is difficult to define. As William Goffman wrote in 1964 [10]:

Relevance is defined as a measure of information conveyed by a document relative to a query . . . the relationship between the document and the query, though necessary, is not sufficient to determine relevance.

Relevance measures how much the "topic" of content is match the request trough the query.

Example: suppose that a user wants information about a specific painting, e.g. Monna Lisa. A page about Picasso is definitively not relevant. A page about Leonardo Da Vinci's works in general, with a mention about Monna Lisa, is appropriate by some means but too generic. Therefore a review on the Monna lisa painting is very relevant. finally the "inner" secret messages of the Monna Lisa paintings (according to theories like the ones of Dan Brown's "Da Vinci code") is again only somehow relevant, because in this case is too specific.

Studying search and traditional information architecture shows that there is no explicit way to show the priority involved with multiple, single topics and their contents. Typically a designer cannot understand how to show the relevance between content acts. Priority in terms of most user interests, most important function, most informative page and so on.

In traditional information architecture the ranking is fixed and is doesn’t have dynamic changes according the variety of needs of users and even the websites owner. Therefore the groups of topics have a fixed sorting or at most multiple sorting options (by date, by title and etc).Honestly theses are not calculated relevance and there are decided in first steps of designing information architecture in terms of designer interests or user needs initial studies which means manually inserted the relevance. Creating a group of topic "masterpieces" or "highlights" out of collections of artworks are the examples try to encounter limitations above.

Stefano Mizzaro and Tefko Saracevic have written histories of relevance, documenting the evolution of this concept among library scientists and information retrieval researchers who have often disagreed about how to measure it [11, 12]. The former tend to take a cognitive, user-centered approach, whereas the latter take a benchmark-oriented approach. This philosophical difference leads to different evaluation approaches, library and information scientists favouring user studies, and information retrieval researchers favoring the use of test collections, particularly those maintained by the Text REtrieval Conference (TREC) [14]. Unfortunately, the TREC approach for measuring relevance has not proven effective for interactive information retrieval systems, and user studies can be prohibitively expensive [15]. Hence, information retrieval researchers accept a definition of relevance as a measure entirely conveyed by a document relative to a query (Goffman notwithstanding). Library and information scientists adhere to a user-centered approach, typically measuring the effectiveness of information seeking support systems through user studies at a task (or higher) level rather than at the query level.

3.3.2 Significance

Significance deals with the importance and therefore currency and reliability of the resource. It can be a query independent measure (i.e., the importance of a page does not depend on your query).

Example: again, a user wants information about Monna Lisa. Let's consider three pages: a page of the louver museum written by an important scholar, a page from wikipedia and finally a page taken from a blog of a tourist that has visited the louver and writes his impressions about Monna Lisa. All these three documents are relevant with respect to the query "Monna lIsa", but... they evidently do not have the same importance! As a user, I expect to find the first tow pages (Louver museum and Wikipedia) ranked in the very first results, and the third page as an additional resource.

Faceted search does not take care of these two issues. The results are relevant, but we cannot sort them by relevance and significance. If a user exploring all the cultural venues in Italy, he wants the Coliseum and the St. Peter Basilica to appear before a small church near Como. If I'm looking only at the Como area, the small church should become important. Consequently one of the solutions could be calculating a ranking measure that estimates relevance and significance from: (1) the metadata attached to items (2) the specific current query.

4 Proposals for solutions

4.1 SEE-IA elements descriptions

The main objective of this thesis is to instruct attendees about the capabilities and advantages of integrating navigation and search for large collections in a seamless, flexible manner that helps users find things quickly and browse items comfortably. The main idea is to take advantage of faceted hierarchical metadata in a systematic way; to specify how to organize this metadata to reflect different user tasks, and how to create a navigation structure that flexibly responds to changes in user's information seeking patterns while at the same time retains consistency in the views of the information structure.

A well-established principle of human memory is that it is often easier to recognize a word or name than it is to think up that word. Thus in many situations it is useful to prompt the searcher with information related to their information need. Browsable information structures, such as links on a Web site or a table of contents for a book, give an overview of the contents of a collection, allowing the searcher to navigate to the information of interest by following links or narrowing by selecting categories. Information structures can also impose an organization on the results of search. To be fully effective, navigation interfaces should allow the user to interleave keyword queries within existing information structures, smoothly integrating navigation with search. This means that after a keyword search, results should be organized into the navigation structure, and that after navigation steps, keyword search should be available over the current subset of information items.

According to [19]In search interfaces, category systems are the main tool for navigating information structures and organizing search results. A category system is a set of meaningful labels organized in such a way as to reflect the concepts relevant to a domain. In search interfaces, categories are typically used either for selecting a subset of documents out from the rest, thus narrowing the results, or for grouping documents, dividing them into (potentially overlapping) subsets, but keeping the documents visible. They can also be used for ordering and sorting search results.

Category system structure in search patterns is usually one of flat, hierarchical, or faceted. A flat list of categories works well for presenting a list of choices with which to narrow the contents of a collection, but needs to be limited to a small set in order to be scannable. Hierarchical (or tree-structured) category systems are useful and can be easy to understand for relatively simple information structures. However, a problem with assigning documents to single categories within a hierarchy is that many information items are best described by multiple different categories simultaneously.

This use of hierarchical faceted metadata provides a usable method for allowing users to browse information collections according to multiple categories simultaneously (Hearst, 2000, Hearst et al., 2002). The main idea is to build a set of category hierarchies, each of which corresponds to a different facet (dimension or feature type) that is relevant to the collection to be navigated. Each facet has a set of labels associated with it, and if this set is large, it may be organized into a hierarchy. After the facet hierarchies are designed, each item in the collection can be assigned any number of labels from any number of facets. In a properly designed faceted navigation interface, the user can browse the information collection from any of the different facets as a starting point, and after starting with one facet, can then navigate using any other facet. Usability results suggest that this kind of interface is highly usable for navigation of information collections with somewhat homogeneous content (English et al., 2001, Hearst et al., 2002, Yee et al., 2003).

According to [18] the aim is to proposes seamlessly integrate features of engineered Information Architecture frameworks (any, of the several proposed) with Search mechanisms.

We need to gain,

- Integrate browsing and searching seamlessly
- Support exploration and learning
- Avoid dead-ends, and “lost ness”

In a traditional Information Architecture, however, the user cannot combine multiple "sibling" classifications, e.g. looking for a detective story novel settled in 1950s. The only way to combine different classifications in this paradigm is by nesting them. But this means forcing designers to make one criterion of classification (often the type of resource: novels, books, etc.) prevail over the others by making it the only starting point for navigation. This choice would imply that users could not get, for example, all the fictional works settled in middle ages, regardless of the type of medium, because they first have to choose the medium.

Based on latest Research from Politecnico di Milano on information architecture which has been deeply focused on communicative aspects .The main idea of integrating search and information architecture proposed is that the interaction between the user and the web application should resemble a structured dialogue, inspired by rhetoric principles, with the purpose of both inform and persuade.

Creating multiple hierarchical schemas may be a solution, but would result in a confusing "offer" to the user – who could not try alternative combinations quickly – and would probably miss desired combinations. In "traditional search engines" although users can find what they are looking for, by means of keywords, this solution show some negative aspects: no communication strength, no branding, no serendipity, and no shop-window would be there. In addition, only users with specific goals (translated into specific sets of keywords) in mind would be satisfied [18]. Users with ill-defined needs or trying to "browse around" would not grasp any clue to start from.

To have optimality to support a seamless user experience, today there are approaches which propose integrated information architecture and search engines with appropriate search paradigms like facet search [21].

Users find information by browsing, searching, asking questions, often employing a combination of these approaches. Information architecture should ideally support the configuration of approaches users employ when finding information. Here the main idea is to create an information architecture which is not static and it has ability to define some elements which can be combined differently. Topics can be used in different aspects of user needs; the may have a leading role and compliance the focus of his interests or having a supporting responsibility like an introductory content to facet. Having this dynamical schema makes the information architecture putting different masks in various user requirements. For example, in a website of paintings we have topics connected to the paintings (the artist, the artistic monument, etc.) the schema gives us simultaneously (1) search for artworks (2) introduce and explain artworks. However we could "pivot" on the artists so now our focus is the artists and we could use artworks properties as a strategy to filter artists; we may want to filter artists that made many portraits.

4.2 Navigating of related information

As already mentioned in previous chapters dealing with related items is one the most crucial parts of the websites form the navigational aspects .Because facing large websites with complex information can contain several related items, understanding how to navigate from different substances and back is a fundamental task that may not be intuitive; especially in situations where you are trying to response several users communication needs. We should propose solutions for all different kinds of browsing and illustrate a many to many relationship by means of navigation.

According to new research paradigms with my colleagues, there are basically three main strategies. In below we are proposing all the navigational methods basically on the appropriate features of the facet search.

4.2.1 Forward navigation

The idea is similar to transition acts but here we can have the navigation in the level of group of topics.

E.g. user goes to browse artists, he filters them by nationality, and then he will have some links to move from those artists (artists from Italy) to their paintings

4.2.2 Inline approach

In this approach there is a possibility to filter the objects with their properties.

E.g. while User is navigating paintings, and he can filter the paintings by artists .the highlight issue is that now we have this opportunity to also flittering the artist by their properties. User selected the subject: landscape, then he decides to filter by a group of artists not a single artist.

We can compare the Forward and Inline navigation in this sense that in the first one, the user is following a chain and moving in a line while n the second user starts from the target, move to something related and finally go back to the target the same as a circle.

4.2.3 Inference of metadata

Here the idea is that the user involves his inferring towards the structures of related metadata. Here the idea is to navigating the parents by the features of their children in hierarchies.

E.g. user accesses from artworks to artists by filtering the nationality .we already have from the forward navigation an "artist country" facet, so in a certain sense, he is "inferring: that a painting is from the artist's nationality since the artist nationality is known.

4.3 Relational browsing

The Semantic Web—as a web of data—can be viewed as a rich, multidimensional hypertext system, on which browsing—as 'regular' hypertext browsing or faceted navigation—can be performed.

'Semantic web' browsing is about exploring the context of semantic data within semantic data, while semantic 'web browsing' is about exploring the semantic context of terms present within a regular web resource and subsequently is about exploring the resource's domain semantics.

The primary purpose of web navigation is to help people to move forward. It is not to tell them where they have been, or where they could have gone. An advantage of Contextual Navigation is that the user may discover the existence of information they would not have otherwise seen because they were buried in the lower down search results. Contextual Navigation requires no manually built taxonomy, configuration, algorithmic training or setup. Once the search index has been built, the navigation is immediately available. The dynamic sub-topic suggestions are derived from the information being indexed by the search engine, so the sub-topics are meaningful in the context of the information being searched and they use the same terminology. These sub-topic suggestions form a hierarchical navigation structure allow the user to click on a suggestion to submit the suggestion as a query to the search engine. By doing this the user can iteratively refine their query, each time being presented with a smaller set of search results to visually scan through. It is also fully customisable user interface for seamless integration.

Examples of these dynamically derived sub-topics can be seen on the left-hand side in the screenshot below.

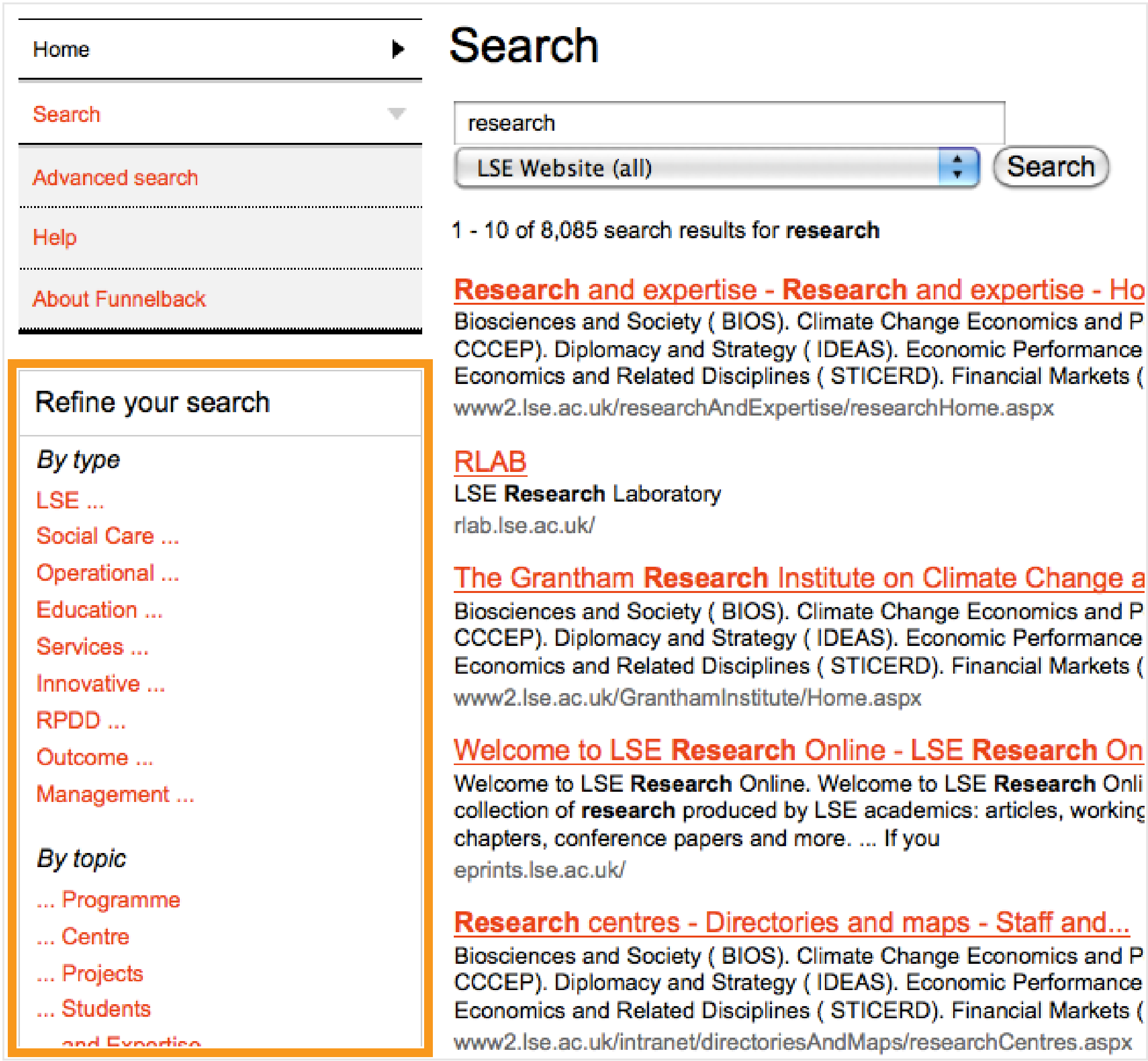


Figure 4-1 example for relational browsing

We can come to these significances for relational browsing 1) Move from a specific topic to other related which reminds us the contextual navigation or transition acts 2) move from a set to another (related) set. The goal of contextual browsing is to help users locate the attributes that are important in specific contexts as quickly and easily as possible. If some attributes are perceived as important in all contexts, then they should be organized accordingly.

There are several managerial implications outcomes for derives from this concept; First, the contextual interface of multi-faceted categorization can be applied to any digital information that needs to be categorized for browsing purposes. For instance, it would be suitable for archives or digital libraries that store large amounts of metadata. The search results for queries should also be categorized and a contextual browsing interface could be adopted by search engine operators to display search results. In particular, contextual multi-faceted categorization could be appropriate for e-commerce websites, where products can be categorized based on various attributes.

Second, the contextual interface can be used as a kind of information seeking aid to help guide users when they browse a large number of facets, instead of browsing at random. Contextually appropriate attributes can thus be chosen quickly and irrelevant features can be screened out. Consequently, users can save cognitive effort when choosing the information they need from an overwhelming number of attributes and, at the same time, maintain the browsing flexibility inherent in a multi-faceted browsing system.

Rather than recommending the results directly to users, the context-sensitive interface provides non-intrusive, indirect guidance during the information seeking process. This gives users more control over the facets and paths they browse. At the same time, our process-orientated approach facilitates effective browsing by embedding a user-defined relevance mechanism in the display of facets on the interface. This kind of browsing interface allows users to control the flow of information. Therefore, it lets users explore new worlds and make serendipitous discoveries while browsing, which a result-oriented approach could not offer.

Finally, the contextual interface can also aid decision-making, especially for consumers who shop on the commercial websites. After browsing the attributes of products, consumers might want to compare the performance of available alternatives in terms of the attributes of interest. Therefore, the contextually relevant attributes can be incorporated into the interface designed for decision-making to help consumers compare the attributes that they perceive as important. To summarize, the contextual interface can be used as a tool for the complete process of information searching, including information categorization, information seeking and decision-making.

5 Conclusion and Future work

Coupling engineered Information Architectures (offering strong organization, powerful navigation, context orientation, and branding and communication strength) with features provided by search patterns and navigation mechanisms can provide at the same time effective access, the proper communication messages (e.g. "what is more important and relevant in the website") and "brand" (e.g. "what is the overall signature of the organization behind the website").

As mentioned before to this end, "neutral" search mechanisms can be ineffective.Faceted navigation provides quick identification of desired information but also a different, playful and highly interactive exploring experience, therefore enhancing both findability and discovery/serendipity.

As seen in several chapters throughout this thesis, an oft-preferred method of search makes use of small steps that allow users to retain the context of their activities. Navigation structures are useful both for imposing an organization on retrieval results, and for re-ordering, narrowing, and filtering those results. An information structure can also provide a useful starting point for getting to know a collection of information.

This work has discussed an extensive range of problems which users dealing with, during the process of browsing and seeking information. Problems in hierarchical structure, group of topics, single topics and contextual navigation. then recalling different search pattern mechanisms It was seen that although hierarchical organization can work for some smaller collections, faceted navigation with query previews has been found to be highly successful for searching within a wide range of information collections. Faceted navigation has the drawback of requiring the category structure to be designed and assigned to items, although automated methods are under development. Most likely in future, hybrids with somewhat more flexible information structures will be explored.

According to particularly focus on the work carried out by the Politecnico di Milano the originality of the novel approach therefore lays in the fact that the overall navigation structure originating from information architecture becomes dynamic by effect of the hybridization with faceted search: access structures are combined according to multiple classifications, adapting to user wishes.

The approach, however, needs further investigation:

- There is space for integrating information architecture and search at deeper level: entire branches (website sections) of the overall hierarchy, or even the overall navigation structure itself could be simplified and made dynamical by means of properly designed hierarchical facets. Obviously this would have a strong impact on how website interface and navigation are conceived.
- If faceted search can used to “simulate” access structures, it could be employed to enhance transition strategies based on semantic relationships.

6 References

[1] Peter Morville, Louis Rosenfeld,Information Architecture for the World Wide Web, Designing Large-scale Web Sites , August 2002

[2] Stefano De Caro, Italian Ministry for Cultural Heritage; Nicoletta Di Blas and Luigi Spagnolo, HOC-Lab, In Search of Novel Ways to Design Large Cultural Web Sites, Politecnico di Milano,Italy

[3] Davide Bolchini, and Paolo Paolini, Interactive Dialogue Model: A Design Technique for Multichannel Applications

[4]Luigi Spagnolo, Search driven information design, December2007

[5] Peter Morville and Jeffery Callender,Search Patterns, January 2010: First Edition

[6] An Interview with Peter Morville, Brad Nunnally on March 29th, 2010

http://johnnyholland.org/2010/03/29/search-patterns-an-interview-with-peter-morville/

[7] Hearst, 2000, Hearst et al., 2002, Yee et al., 2003, Hearst, 2006b Research on the Flamenco project, Hearst et al.

[8] Neal, M. Parametric search: evolving information retrieval for the web, 1997. http://www .intranetjournal.com/ features/cadis-1.shtml.

[9] Daniel Tunkelang ,Faceted Search , 2009,www.morganclaypool.com

[10]Goffman, W. On relevance as a measure: Information Storage and Retrieval, 1964.

[11] Mizzaro, S. Relevance: the whole history. Journal of the American Society for Information Science 48(9): 1997 (September).

[12] Saracevic, T. Relevance: a review of the literature and a framework for thinking on the notion in information science. Part II. Advances in Librarianship 2006.

[13] Text retrieval conference http://trec.nist.gov/.

[14] Voorhees, E. On test collections for adaptive information retrieval. Information Processing and Management 2008 (November).

[15] Stefano De Caro, Nicoletta Di Blas, Paolo Paolini, Luigi Spagnolo, Search-enhanced web information architecture for findability and discovery of archeological heritage: the MiBAC-DGA case study,

[16] Artist rising website, http://www.artistrising.com/

[17] http://www.freebase.com/view/visual_art/views/artwork

[18] Luigi Spagnolo, Paolo Paolini, Davide Bolchini, Nicoletta Di Blas, Beyond Findability ,Search-Enhanced Information Architecture for Content-Intensive Rich Internet Applications

[19] Marti A. Hearst, Search User Interfaces, Cambridge University Press 2009

[20] Morville P., Ambient findability: what we find changes who we become, O'Reilly Media, Inc. , 2005

[21] Quintarelli E., Resmini A., Rosati L. Facetag: Integrating Bottom-up and Top-down. Classification in a Social Tagging System, , 2007

[22] Crystal, Abe. Facets Are Fundamental: Rethinking Information Architecture Frameworks, Technical COMMUNICATION • Volume 54, Number 1, February 2007

[23] Brinck, T., D. Gergle, and S. Wood. Usability for the Web: Designing Web sites that work. 2002.