

Facetag: Integrating Bottom-up and Top-down Classification in a Social Tagging System

Emanuele Quintarelli

Reed Business Information, Milano (Italy). info@infospaces.it

Andrea Resmini

CIRSFID, University of Bologna, Bologna (Italy). root@resmini.net

Luca Rosati

University for Foreigners, Perugia (Italy). luca@lucarosati.it

Abstract — FaceTag is a working prototype of a semantic collaborative tagging tool conceived for bookmarking information architecture resources. It aims to show how the widespread homogeneous and flat keywords' space created by users while tagging can be effectively mixed with a richer faceted classification scheme to improve the “information scent” and “berrypicking” capabilities of the system. The additional semantic structure is aggregated both implicitly observing user behaviour and explicitly introducing a compelling user experience that facilitates the end-user creation of relationships between tags. To evaluate the real benefits introduced by this approach, preliminary user research on the initial prototype is outlined. Extensive evaluation and results will be discussed in a future paper.

1. INTRODUCTION

Collaborative tagging systems have been largely adopted by end-users as useful and powerful tools to organize, browse and publicly share personal collections of resources on the World Wide Web through the introduction of simple metadata.

The aggregation of user metadata is often referred to as a folksonomy, a user-generated classification, emerging through bottom-up consensus while users assign free form keywords to online resources for personal or social benefit. Del.icio.us <<http://del.icio.us/>>, Flickr <<http://www.flickr.com/>>, 43things <<http://www.43things.com/>>, Furl <<http://www.furl.net/>> and Technorati <<http://www.technorati.com/>> are web-based collaborative systems for building shared databases of items, enriched by a flat metadata vocabulary that can be used to perform metadata-driven queries, to monitor change in areas of interest or to discover emergent trends, such as the hottest / most popular topics in the system (Quintarelli 2005).

In the past, folksonomies have often been seen as orthogonal to taxonomies and controlled vocabularies: the latter rigid, hierarchical and organically hand-crafted by professionals a priori; the former flat, inclusive and emerging from bottom-up users' input and consensus (Quintarelli 2005). In a flat tagging system each document can be retrieved through a simple set of keywords, collaboratively introduced by users to describe and categorize the document, very much like in a keyword-based search process in which descriptive terms can be used to get a set of applicable items.

Despite their low cognitive cost, their capability of matching users' real needs and language and their great value in a serendipity research task, folksonomies imply however a lack of precision, a very low findability quotient (especially in a known-item approach) and a limited scalability for the intrinsic variability of language (Quintarelli 2005).

As a result of the inherently inconsistent, evolving and much variable process of associating words and meanings, tagging systems are also implicitly plagued by a number of linguistic issues which include polysemy, homonymy, plurals, synonymy and basic level variation which do not appear

easy to solve (Golder & Huberman 2005). Any of these problems can dramatically reduce the effectiveness of the application, mining the benefits brought on by the use of tagging systems.

In addition, tags have recently started to be used by bloggers as reading-aids to help users identify articles and posts of interest, providing as such a complimentary structure over a purely chronological list of text pieces. This approach marks a major shift, in that tagging also becomes a tool to maximize findability and browsability without limiting the reader to only access the most popular or recent tags as in common tag clouds (Feinstein & Smadja 2006).

Tag clouds are widely used visual interfaces for information retrieval that provide a global contextual view of tags assigned to resources in the system. In such a structure, the most popular tags are usually displayed through an alphabetically ordered list with the font size increasing with the tag's relevance. Users browse the cloud, scanning hyperlinks to recognize information of interest (Hassan-Montero & Herrero-Solana 2006).

Flat tag clouds as currently implemented are not sufficient to provide a semantic, rich and multidimensional browsing experience over large tagging spaces. There are several reasons for this:

1. Choosing tags by frequency of use inevitably causes a high semantic density with very few well-known and stable topics dominating the scene (as seen on RawSugar, <<http://www.rawsugar.com/>>);
2. Providing only an alphabetical criterion to sort tags heavily limits the ability to quickly navigate, scan and extract, and hence build a coherent mental model out of tags;
3. A flat tag cloud cannot visually support semantic relationships between tags. We suggest that these relationships are needed to improve the user experience and general usefulness of the system;
4. Current tag clouds often fail to provide complex logical operation over tags. Simply clicking on a tag is not enough to enable a smooth and powerful exploration or refinement.

Even if FaceTag doesn't promise to address all of these issues, we believe our approach can limit the impact of linguistic complications such as polysemy, homonymy and basic level variation while introducing an innovative, multidimensional and more semantic paradigm for organizing, navigating and searching large information spaces through tags.

To reach this goal, FaceTag contributes to social tagging systems in three ways:

1. The use of (optional) tag hierarchies. Users have the possibility to organize their resources by means of parent-child relationships;
2. Tag hierarchies are semantically assigned to editorially established facets that can be later leveraged on to flexibly navigate the resource domain;
3. Tagging and searching can be mixed to maximize findability, browsability and user-discovery.

2.RELATED WORKS

The widespread adoption of tagging systems by end-users has greatly stimulated discussion about their long-term implications inside the information studies community and the so-called blogosphere.

Initial groundwork for the creation of metadata ecologies made by the simultaneous use of different classification approaches has been provided by Morville, Campbell and Fast.

In *Ambient Findability*, Morville (Morville 2005) states that tagging has its own proper place inside information architecture theory and practice, suggesting that these systems can be

productively considered a complementary fast moving layer over slower layers represented by more traditional information architecture practices.

Grant Campbell and Karl Fast reinforced this position showing how pace layering theory can be combined with the complexity and resilience theories to provide a working model of interaction between folksonomies and conventional information architecture (Campbell & Fast 2006).

A number of examples have inspired the idea of leveraging facets and improved our understanding of their relationship with clusters and taxonomies.

A useful foundation for the comprehension of the way users choose tags in a social space has been provided by Golder and Huberman (Golder & Huberman 2005). They have analyzed the structure and dynamics of collaborative tagging systems trying to extract stable patterns and recurrent tag types in del.icio.us <<http://del.icio.us/>>. Their work has effectively synthesized the principal issues affecting flat tagging spaces and Facetag aims to address some of these issues .

To reach this goal, our work has followed the path laid out by other publications and sites.

Concrete examples of tagging ecologies have been provided by Mefedia, Etsy, LibraryThing and Rawsugar.

With Mefedia <<http://mefedia.com/>>, Dutch Information Architect Peter Van Dijck was among the first to propose structured tagging, mixing facets and tags. In this case, tags are statically (i.e. manually) assigned to a fixed number of editorially designed facets. This editorial effort on users' provided tags makes large tagclouds more significant and single tags more meaningful.

The social bookmarking site RawSugar <<http://www.rawsugar.com/>> was the first one to introduce algorithmically generated hierarchies of tags. The end result is a forest of tag trees that can be used to support wayfinding both in search and navigation.

Etsy <<http://www.etsy.com>> is a niche marketplace for handmade things where people can sell and buy items. The site provides a complete set of different access dimensions (implicitly raw Ranganathan facets for space, time, material, topic, colors, owners, etc) to their information through flashy visualization clues. Etsy is also mixing categories (a flat topic taxonomy) with user generated tags (used as sub categories).

Finally, LibraryThing <www.librarything.com> is a social site that let users catalogue and tag books. Here subject headings are shown side by side with people assigned tags. Users can browse using a top down approach or surf horizontally leveraging tags.

On a more theoretical perspective, a discussion of faceted classification models in current web sites and the relationship between facets and tagging has been presented by Travis Wilson at IA Summit 2006 (Wilson 2006).

Marti Hearst and The Flamenco project investigated for thirteen years how faceted interfaces can help users flexibly navigate and search through large information spaces (Hearst, The Flamenco Search Interface Project). Their studies and papers are easily recognizable in our approach (see for example Hearst 2006a) especially affecting the interface design.

Finally, Sam H. Kome, Heymann, Garcia-Molina and Bar-Ilan have recently proposed possible ways and benefits regarding the introduction or elicitation of relationships already contained inside folksonomies.

Hierarchical relationships can be implicitly found in tagging systems as showed by Sam H. Kome (Kome 2006), while Heymann and Garcia-Molina presented a simple algorithm to automatically convert tags associated to objects into a hierarchical taxonomy (Heymann & Garcia-Molina 2006).

In conclusion, Bar-Ilan (Bar-Ilan et al. 2006) compares unstructured (freely assigned tags) and structured tagging (tags assigned to predefined metadata elements), suggesting that structured tagging may be able to produce stronger user guidance, hence possibly resulting in higher quality descriptions.

3.OVERVIEW OF SEMANTICAL STRUCTURES IN TAGGING SYSTEMS

Usability studies show how information seekers in large domains of objects prefer meaningful groupings of related items, in order to quickly understand relationships and so decide how to proceed (Hearst 2006a). In other words, it seems quite clear that without any means to explore and make sense of large quantities of similar items, users feel lost and fail to complete their information seeking tasks.

How to generate and navigate such groups from a flat set of objects is anyway a different matter. Both clustering and faceted classification have been proposed in the past as useful techniques which allow searchers to easily browse and navigate information spaces.

Clusters

Document clustering refers to the act of grouping of items according to some measure of similarity, typically searching for identifiable repetitive patterns of words and phrases.

Some advanced tagging systems like Rawsugar and Flickr are already using clusters to address the issues that plagued the first generation of folksonomy-based applications: clusters help reduce the semantic density and improve the visual consistency of tag clouds. Moreover, clustering is automatable, can be used to refine vague queries and to disambiguate ambiguous search keywords.

Nonetheless, clustering techniques and algorithms are not perfect and often generate messy groups which are generally hard to predict. These groups also tend to conflate many different dimensions becoming also hard to label in ways that are meaningful for users. Moreover, clustering does not generally allow issuing refinement and follow-up queries, thus heavily limiting the explorative capabilities of the system.

For these reasons, usability results show that users prefer clear hierarchies with categories at uniform levels of granularity over the messy, unpredictable and unlabeled groupings typical of clustering techniques (Hearst 2006a).

Hierarchical Facets

At the other end of the classification line, traditional hierarchical categories are coherent and complete systems of meaningful labels which systematically organize a domain. The main drawback of this approach is that a single a priori and monolithic hierarchical organization rarely has the capability to match the varied ways of thinking and organizing the world of different users.

Hierarchical faceted metadata has shown to be a promising middle ground, able to satisfy the needs of a wide range of users with different mental models and vocabularies (Yee et al. 2003).

Facets are orthogonal descriptors (i.e. categories) within a metadata system. Each facet has a name and it addresses a different conceptual dimension or feature type relevant to the collection. Facets can be flat or hierarchical and they can be assigned single or multiple values. Thus a faceted search interface requires that each object in the collection is classified through labels from different facets.

In a hierarchical faceted navigation tool, choosing a label from one of the facets is equivalent to performing a disjunction over all the labels beneath the selected one, while choosing labels from different hierarchies builds a query that is a conjunction of disjunctions over the selected labels and their sublabels. In this kind of interface, users can navigate multiple faceted hierarchies at the same time (English et al. 2002b). Usability studies show how this approach is preferred over single hierarchies because users feel in control without getting lost (English et al. 2002b, Yee et al. 2003).

Additional features exposed by faceted based interfaces are the suggestion of logical alternatives at each navigation step and avoidance of dead ends.

For these reasons, faceted metadata can be used to support navigation along several dimensions simultaneously, allowing seamless integration between browsing and free text searching and an easy alternation between refining (*zooming in*) and broadening (*zooming out*) the query while retaining a feeling of control and understanding (English et al. 2002b). The major benefits resulting from this approach include: a strong reduction of the mental work, favoring recognition over recall and better support for exploration, discovery and iterative query refinement (Hearst 2006a).

Again, usability studies attest how hierarchical faceted interfaces are preferred over simpler keyword based search interfaces and how they can be easily understood by the average user (Yee et al. 2003) if iteratively designed and tested to address usability issues (English et al. 2002a).

4.OVERVIEW OF FACETAG

Until today, one of the main limitations of hierarchical faceted categories was the lack of a good automated process for both creating the categories and associating items to the hierarchy of labels under each facet (Hearst 2006a).

We decided to avoid the issue entirely and use no algorithmic round-ups: FaceTag is built around the notion that the users provide the structure. It especially aims to investigate how a hierarchical and faceted metadata structure can be added to user generated content facilitating use of tags provided by end users in collaborative systems. Barriers to effectively using FaceTag are limited through a careful user interface design.

Faceted analysis: the faceted scheme construction

Although *facet* and *faceted* have become very common terms in the information architecture field, their application often falls far from its original meaning. The attribute faceted, indeed, is used in a large variety of meanings, and is often referred loosely to the availability of means to search by different keys (La Barre 2004). The full theory of faceted classification, as it has been developed by Ranganathan and the Classification Research Group (CRG) and which includes rules for citation order and notation, is less widespread as a backend for website organization; remarkable exceptions are offered by projects staffing librarians, such as FATKS (Slavic 2002).

So, we thought to apply faceted classification to resources in the IA field itself adhering to the principles of the original library theory, in order to leverage on its potentialities and obtain maximum benefits. In such perspective, our design was inspired by these projects:

1. the Flamenco project <<http://flamenco.berkeley.edu/>>
2. Facetious <<http://demo.siderean.com/facetious/facetious.jsp>>
3. Etsy <<http://www.etsy.com>>¹

The choice of facets is based on the CRG theory (Vickery 1960) and, although we have remained faithful to the CRG standard categories (Broughton 2001), we have reviewed the resulting facets to fit into a more semantic perspective. Indeed, an aspect often underestimated on the World Wide Web is that both Ranganathan and the CRG described a generic schema for faceted classification, which every actual schema can refer to. Thus, in a faceted classification project one does not have to rebuild the schema from scratch every time, but may follow a constant guideline while building one's main categories (i.e. facets). CRG postulates 11-13 general categories. In the table below we show the matching between CRG standard categories and IA-related categories that were used to define our facets. Labels are yet to be formalized and facets will be verified in future works. It is our believe that Ranganathan's list and the CRG postulates should be considered as guidelines and not as a restrictive framework. User needs are paramount and should be used to determine the most valuable facets for every system (See #9).

¹ Both Facetious and Etsy mix proper facets and metadata (formal proprieties of an item).

Table 1: FaceTag facets definition by CRG standard categories.

CRG	FaceTag
Thing	[Documents, resources]
Type	Resource Types (e.g. online report, case study...)
Part	--
Property	Language
Material	Themes
Process	--
Operation	[Activities] ²
Product	Deliverables
Byproduct	--
Patient	Purposes, Markets (e.g. Industry, Health ...)
Agent	People
Space	[Country]
Time	Date

A preliminary analysis of a corpus of IA resources from the Information Architecture Institute Library <<http://iainstitute.org/library/>> allowed us to define six facets which appeared to be suitable for the classification of IA resources.

Table 2: FaceTag facets and examples of foci

Facet	Samples
Resource Types	white paper, case study, blog > enterprise web,
Language	<i>predefined values (based on ISO Standard ISO 639-2)</i>
Themes	competitive analysis, classification > facets, web 2.0 > folksonomies, information design > navigation design > breadcrumbs
People	weinberger, e. reiss, morville
Purposes	industry, public administration, health, software > companies > google, education > conferences > www2006

²Facets in brackets have been considered of secondary importance and discarded.

Facet	Samples
Date	<i>date of publication (not insertion)</i>

The foci listed near some of the facets serve the only purpose of making the facets self-explanatory. In the actual implementation, since tags are our foci, foci will be user-generated, with the only exception of the language facet, which will use a predefined list of languages in the ISO 639-2 notation, and the date facet, which will receive a software-generated timestamp upon resource creation.

5.USING FACETAG

FaceTag deals with tags, facets, resources and users in two distinct ways: a browsing / searching mode and a bookmarking / editing mode. These are two different activities, to which the user interface adapts providing different aiding tools (navigation, resource and user management) and different behaviours (zooming, tag autocompletion and tag suggestion) respectively. The browsing interface is used to look for resources; the bookmarking interface can be only accessed by authenticated users and it is used to add new resources to the system. It also provides tools to administer one's profile.

The screenshot shows the FaceTag homepage. At the top, there's a navigation bar with links: Home, Take a tour!, Get in touch!, and About FT. Below this is a banner stating: "FaceTag is a working prototype of a semantic collaborative tagging tool conceived for bookmarking Information Architecture resources. Use it to save, find and manage your bookmarks or to discover new interesting connections."

The main content area features five facets:

- Language:** A dropdown menu currently set to "all".
- Resource type (24):** A list including article (7), blog + (3), magazine + (5), paper (4), posters (1), toolkit (2), and tutorial (2).
- Themes (33):** A list including contextual inquiry (1), deliverables + (2), design (1), diagrams (1), ethnography (1), evaluation (1), folksonomies (4), information architecture (6), interface design (2), intranets (4), knowledge management (2), navigation design (1), scent of information (1), social classification (2), tag (1), tagging (2), and tags (1). A "more tags" link is present.
- People (11):** A list including david weinberger (1), quintarelli (2), isko + (1), jared spool (1), rosati (1), shiv.singh (1), step.two.design (1), and stephen.turbek (2).
- Purposes (17):** A list including conference (1), deliverable (1), education + (1), interface design (1), intranet design (4), myproject (1), project planning (1), prototyping (1), research (3), user experience (2), and wireframes (1).

Below the facets is a "Search resources" section with a search bar and a dropdown menu showing "Inf", "Informatics", "Information architecture", and "Information design".

The "Recent bookmarks (85 items)" section is displayed, with sorting options: "order by insertion (newest / oldest)" and "order alphabetically (ascending / descending)". It lists several bookmarked resources:

- Folksonomies: power to the peoples** by andrea at 23 may 2007 20:30 in: article, folksonomies, information architecture, social classification, tag, tagging, quintarelli. URL: <http://www.infospaces.it/>
- A List Apart: Articles: Power To The People: Relative Font Sizes** by emanuele at 23 may 2007 20:30 in: article, magazine > a list apart, interface design, typography, bojan mihelac. URL: <http://www.alistapart.com/>
- Real Wireframes Get Real Real Results - Boxes and Arrows: The design behind the design** by Luca at 23 may 2007 20:30 in: article, information architecture, deliverables > wireframes, stephen turbek, myproject. URL: <http://www.bboxesandarrows.com/>
- GUUUU! - Navigation blindness** by emanuele at 23 may 2007 20:30 in: blog > guuu!, information architecture, navigation design, usability, design. URL: <http://www.guuui.com/>
- KMWorld.com: The BBC's low-tech KM** by Luca at 23 may 2007 20:30 in: magazine, article, knowledge management, intranet, david weinberger, intranet design. URL: <http://www.kmworld.com/>
- Technology Review: Emerging Technologies and their Impact** by Luca at 23 may 2007 20:30 in: magazine, magazine > technology review, user research, contextual inquiry, ethnography. URL: <http://www.kmworld.com/>

At the bottom, there is a "Read more >>" link.

Figure 1: FaceTag homepage

Figure 1 demonstrates the default or browsing mode of FaceTag. This is the screen a user first accesses prior to login. At the top of the current interface is the header area which contains the main navigation tabs, the facets *Resource Types*, *Themes*, *People* and *Purposes*, and the resources area. In default mode, the resource area lists the most recently added bookmarks

The left-most container in the facet area presents filters for *Language* and *Publication date*, and the search box. Language and Publication date are actual facets, but are primarily used as simple filtering tools because of their special, flat nature³.

The remaining facet containers present, left to right, the most used tags for *Resource types*, *Themes*, *People* and *Purposes*. Each facet lives in its color-coded space, and query previews for tags are provided⁴. First-level tags in a hierarchy are followed by a + sign, which can be clicked to expand them and access the complete tree of children tags.

Suppose a user wants to look for a specific subject. She's on the homepage and she starts typing 'Inf' into the search entry field. This is an autocompletion widget, so as soon as she enters the third character, FaceTag starts to suggest possible choices, reading from existing tag data stored in the database: this time FaceTag suggests *Informatics*, *Information architecture*, *Information design*.

She decides that *Information architecture* is what she wants to look for, so she selects that, confirms and clicks *Search*.

The screenshot shows the FaceTag interface. At the top, there's a header with the FaceTag logo and navigation links: Home, Take a tour!, Get in touch!, and About FT. Below the header, a banner states: "FaceTag is a working prototype of a semantic collaborative tagging tool conceived for bookmarking Information Architecture resources. Use it to save, find and manage your bookmarks or to discover new interesting connections." The main navigation area consists of several facets: Language (set to 'all'), Publication date (with 'from' and 'to' date pickers), a Search resources box, and five colored tabs: Resource type (4), Themes (16), People (5), and Purposes (7). The Resource type facet shows 'article (3), blog (1)'. The Themes facet lists various tags like 'deliverables + (2), design (1), diagrams (1), folksonomies (1), intranets (1), navigation design (1), scent of information (1), social classification (2), tag (1), tagging (2), usability (2), navigation (1)'. The People facet lists 'quintarelli (2), jared spool (1), stephen turbek (2)'. The Purposes facet lists 'conference (1), deliverable (1), interface design (1), intranet design (1), myproject (1), research (2)'. Below the facets, a search bar contains the text "Information architecture" with a red 'X' icon. The main content area displays "5 bookmarks" with sorting options: "order by insertion (newest / oldest)" and "order alphabetically (ascending / descending)". The bookmarks are listed with their titles, authors, dates, and tags. The first bookmark is "Folksonomies: power to the peoples" by andrea, dated 23 may 2007 20:30, with tags: article, folksonomies, information architecture, social classification, tag, tagging, quintarelli. The second is "Real Wireframes Get Real Real Results - Boxes and Arrows: The design behind the design" by Luca, dated 23 may 2007 20:30, with tags: article, information architecture, deliverables > wireframes, stephen turbek, myproject. The third is "GUUUI - Navigation blindness" by emanuele, dated 23 may 2007 20:30, with tags: blog > guuui, information architecture, navigation design, usability, design. The fourth is "Intranet Portals and Scent are Made for Each Other" by emanuele, dated 23 may 2007 20:30, with tags: article, information architecture, usability, intranet, scent of information, jared spool. The fifth is "The Lazy IA's Guide to Making Sitemaps - Boxes and Arrows: The design behind the design" by andrea, dated 23 may 2007 20:30, with tags: blog, information architecture, deliverables > sitemaps, diagrams, stephen turbek.

Figure 2: Bookmarks tagged with *Information architecture*

³For example, they cannot currently be part of a tag hierarchy.

⁴Query previews display the number of resources associated with each tag. This is automatically gathered from information stored in the database.

FaceTag *engages* the term and both the facet containers and the resource area adjust: this is called *zooming in* (Figure 2). What the user sees is now a filtered view, a subset of all available resources based on the current active selection. In our example, the returned result set is small as only a short list of bookmarks have been tagged *Information Architecture*.

Actual result sets may be orders of magnitude larger, so she decides she needs to refine her search. She looks at the facets and clicks on *Article* from the *Resource types* facet. FaceTag again engages the term and both the facet containers and the resource area adjust (Figure 3): as we can see, *Resource types* is no longer useful to refine this search and it provides no further possible selections, while the other facets list a smaller number of tags and smaller query preview figures. The resource area is now showing fewer bookmarks.

The screenshot shows the FaceTag web application interface. At the top, there's a navigation bar with links: Home, Take a tour!, Get in touch!, and About FT. Below this is a banner stating: "FaceTag is a working prototype of a semantic collaborative tagging tool conceived for bookmarking Information Architecture resources. Use it to save, find and manage your bookmarks or to discover new interesting connections."

The main interface features several facets on the left and top, and a list of search results on the right. The facets include:

- Language:** A dropdown menu set to "all".
- Publication date:** Fields for "from" and "to" with year pickers.
- Search resources:** A text input field with "tags or keywords..." and a "search" button.
- Resource type:** A facet showing "article" with a red "X" indicating it's the active filter.
- Themes (11):** A list of tags including "deliverables + (2)", "folksonomies (1)", "intranets (1)", "navigation design (1)", "scent of information (1)", "tag (1)", and "tagging (2)", "usability (2)".
- People (3):** A list of names including "quintarelli (1)", "jared spool (1)", and "stephen turbek (1)".
- Purposes (5):** A list of purposes including "deliverable (1)", "interface design (1)", "intranet design (1)", "myproject (1)", and "navigation (1)".

Below the facets, there's a section titled "3 bookmarks" with sorting options: "order by insertion (newest / oldest)" and "order alphabetically (ascending / descending)". The results list three items:

- Folksonomies: power to the peoples**
Folksonomies are web-based collaborative systems for building shared databases of items, enriched by a flat metadata vocabulary that can be used to perform metadata-driven queries.
posted by [andrea](#) at 23 may 2007 20:30 in: [article](#) [folksonomies](#) [information architecture](#) [social classification](#) [tag](#) [tagging](#) [quintarelli](#)
<http://www.infospaces.it/> - [cached](#) - [mail it](#) - [blog this](#)
- Real Wireframes Get Real Real Results - Boxes and Arrows: The design behind the design**
posted by [Luca](#) at 23 may 2007 20:30 in: [article](#) [information architecture](#) [deliverables](#) [wireframes](#) [stephen turbek](#) [myproject](#)
<http://www.bboxesandarrows.com/> - [cached](#) - [mail it](#) - [blog this](#)
- Intranet Portals and Scent are Made for Each Other**
posted by [emanuele](#) at 23 may 2007 20:30 in: [article](#) [information architecture](#) [usability](#) [intranet](#) [scent of information](#) [jared spool](#)
<http://www.uie.com/> - [cached](#) - [mail it](#) - [blog this](#)

Figure 3: Refining the results: looking for *articles* on *Information architecture*

Again, she is not satisfied and clicks *Folksonomies* from the *Themes* facet (Figure 4). Once more, FaceTag engages the term. The facet containers and the resource area adjust themselves: this is the final step for this search, since the result set consists of a single resource that complies with our engaged tags and search terms. As Figure 4 demonstrates, the bookmark *Folksonomies: power to the people* has been tagged with *Information architecture*, *article*, *folksonomies*.

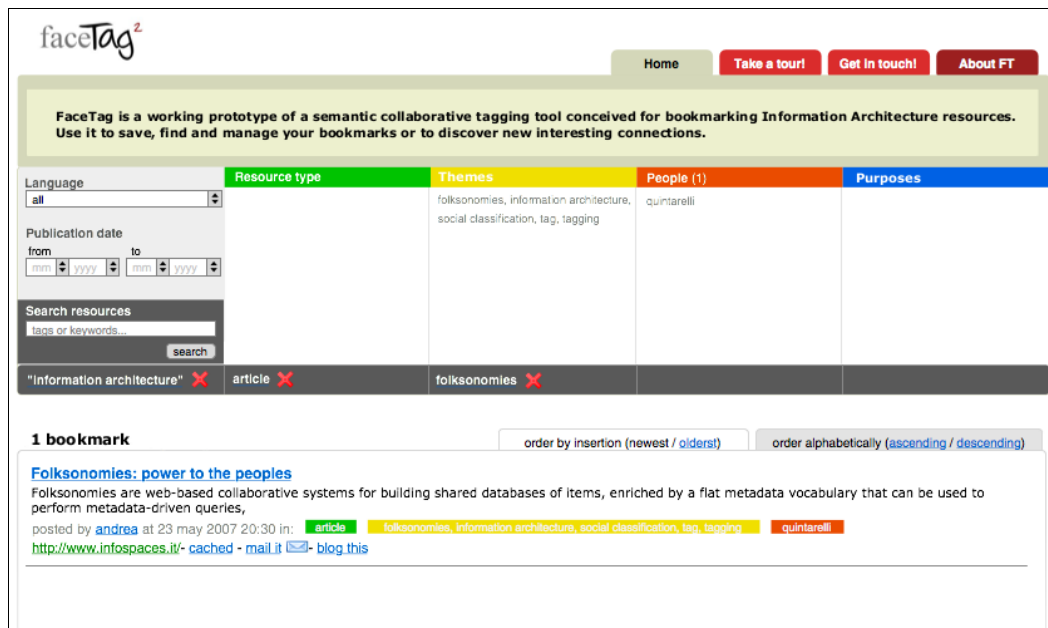


Figure 4: The final result set

Note that the facet *Purposes* in Figure 4 provides no further possible selections. The facets *Themes* and *People* do not list links, but simple grayed-out tags: these are the tags pertaining to the *zoomed-in* resource.

It's worth mentioning that searching and browsing in FaceTag are one seamless process and users can use either way to proceed, mixing them as they deem appropriate. The engage system keeps track of every step taken during a search and engaged tags are listed just below their pertaining facet, in the facet area. Users can disengage every tag they choose any time, in no particular order: this feature set bootstraps 'berrypicking' search strategies and allows users to 'follow the information scent'. A user may nonetheless instead decide to deselect all tags at once and start anew.

If the user logs in, then she can add new resources (Figure 5). The *New bookmark* page requires the user to enter a title for the bookmark, its URL⁵, a rich description using a WYSIWYG embedded editor and any number of tags or hierarchical tags, one facet a time. As can be seen in Figure 5, an autocompletion widget helps the user by suggesting tags and tag placement within existing hierarchies.

⁵These are going to be added automatically if FaceTag is accessed from within a bookmarklet

Facets Evaluation

As anticipated, inside the Faceted Analysis section, preliminary facets were derived from the facets suggested by the CRG schema by comparing these to our current scenario (bookmarks for information architects).

These preliminary facets will be revised through an iterative bottom-up procedure to elicit the (possible and more popular) facets from a wide set of IA bookmarks already online. To this goal we are collecting samples of IA related bookmarks from the IAI Library and from Del.icio.us in order to perform iterative card sorting tests, with different user groups. The purpose of such tests is to figure out the several mental models by which users represent the IA knowledge domain. These results, combined with an analysis of IA-related tags in selected social tagging spaces (such as Del.icio.us, Technorati, Magnolia or Flickr), will provide the basis to tune the facets architecture.

User Interface Evaluation

FaceTag is a social tagging application that combines multiple classification schemes and navigation approaches in a single organic interface.

Providing an intuitive, easy to learn and easy to use interface is probably the single most effective way to support and stimulate the introduction of tags and relationships between them and facets in a semantic tagging application.

The new user interface has been designed through documented heuristics and patterns and verified at each iterative step by small usability tests. More extensive user research will involve the use of think-aloud protocol sessions with more than 5 testers (for each session). Scenarios will include storing bookmarks and retrieving them.

Looking at preliminary results, a critical task addressed by the application is the assignment of new bookmarks and the association of tags to relevant facets. The current interface is rather simple but a number of alternatives leveraging advanced tag suggestion and tag/facet association is under evaluation.

7.CONCLUSIONS

By providing the user with facets to which hierarchical sets of tags relate and pertain and a usable interface which adapts to the ongoing query, FaceTag may solve, through contextualization and user-added semantic value, most of the basic issues connected with polysemy, homonymy and base level variations.

While further testing and usability studies are needed to verify to which extent users are motivated to use our prototype and to introduce structure in addition to flat tags, preliminary user evaluations show how the introduction of hierarchies and facets can improve and disambiguate the meaning of tags giving them a stronger context and a more coherent organization.

Improving on current features, FaceTag aims to provide an advanced tagging experience through other innovative tools or widgets, like a Firefox plugin to seamlessly add new bookmarks while browsing, a WYSIWYG editor to offer drag and drop inclusion of texts and pictures from the web page the user is bookmarking and a history of all the times a bookmark has been tagged.

Future works include publishing a mature release of FaceTag and conducting extensive tests with Information architects.

8.ACKNOWLEDGEMENTS

Special thanks go to Luca Mascaro and Diego La Monica. Luca has been in charge of designing and testing the new user interface: his advice has deeply influenced the ongoing visual solutions and the global look and feel of FaceTag. Diego is responsible for large amounts of that very code

that allows us to make that look and feel work in the real world and for straightening out old algorithms while designing new better ones.

Finally, we would like to thank for their editorial assistance Kathryn La Barre (Assistant Professor at University of Illinois at Urbana-Champaign) and David Swedlow (Small World Labs in Austin TX). Both helped us frame our vision of folksonomies and Kathryn provided precious insights into facet theory.

References

- Bar-Ilan J., Shoham S., Idan A., Miller Y., Shachak A., (2006) *Structured vs. unstructured tagging – A case study*, WWW2006, Edimburg <<http://www.rawsugar.com/www2006/12.pdf>>.
- Broughton, V. (2001) *Klasifikacija za 21. stoljeće: nacela i struktura Blissove bibliografske klasifikacije* (= *A classification for the 21st century: principles and structure of the Bliss bibliographic classification*), Vjesnik bibliotekara Hrvatske, 44, 1-4, p. 38-51; trad. it. Una classificazione per il 21° secolo: principi e struttura della Classificazione bibliografica Bliss, AIB-WEB. Contributi, <<http://www.aib.it/aib/contr/broughton1.htm>>.
- Campbell, G.D., Fast, K.V., (2006) *From Pace Layering to Resilience Theory: The Complex Implications of Tagging from Information Architecture*, Proceedings of IA Summit 2006 (Vancouver, March 23-27, 2006), ASIS&T <http://www.iasummit.org/2006/files/164_Presentation_Desc.pdf>.
- Chi, E.H. - Pirolli, P., Chen, K. – Pitkow, J. (2001) *Using Information Scent to Model User Information Needs and Actions on the Web*, Proceedings of the SIGCHI conference on Human factors in computing systems (Seattle, Washington, 2001), ACM Press <<http://www2.parc.com/istl/projects/uir/publications/items/UIR-2001-07-Chi-CHI2001-InfoScentModel.pdf>>.
- English, J., Hearst, M., Sinha, R., Swearingen K., and Yee, P., (2002a) *Hierarchical Faceted Metadata in Site Search Interfaces*, CHI 2002 Conference Companion <http://flamenco.berkeley.edu/papers/chi02_short_paper.pdf>.
- (2002b) *Flexible search and browsing using faceted metadata*, Unpublished Manuscript <<http://flamenco.berkeley.edu/papers/flamenco02.pdf>>.
- Feinstein, D., Smadja F., (2006) *Hierarchical Tags and Faceted Search. The RawSugar Approach*, Proceedings of SIGIR 2006 (August 6-11, 2006, Seattle, Washington).
- Flamenco Group (2002) *How to Build a Flamenco instance* <<http://bailando.sims.berkeley.edu/flamenco/howtobuild/howtobuild.html>>.
- Gnoli, C., Marino, V., Rosati, L., (2006) *Organizzare la conoscenza. Dalle biblioteche all'architettura dell'informazione per il Web* (= *Organizing Knowledge. From Libraries to Information Architecture for the Web*), Tecniche Nuove.
- Golder, A.S., Huberman, B.A., (2005) *The Structure of Collaborative Tagging Systems*, Information Dynamics Lab <<http://arxiv.org/pdf/cs.DL/0508082>>.
- Hassan-Montero, Y., and Herrero-Solana, V., (2006) *Improving Tag-Clouds as Visual Information Retrieval Interfaces*, International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006 <http://www.nosolousabilidad.com/hassan/improving_tagclouds.pdf>.

- Hearst, M.A. (2006a) *Clustering versus faceted categories for information exploration*. Communication of the ACM April Vol 49, No.4
<<http://flamenco.berkeley.edu/papers/cacm06.pdf>>.
- (2006b) *Design Recommendations for Hierarchical Faceted Search Interfaces*, ACM SIGIR Workshop on Faceted Search <<http://flamenco.berkeley.edu/papers/faceted-workshop06.pdf>>.
- *The Flamenco Search Interface Project* <<http://flamenco.berkeley.edu/pubs.html>>.
- Heymann, P., Garcia-Molina, H., (2006) *Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems*, Technical Report InfoLab
<<http://dbpubs.stanford.edu/pub/2006-10>>.
- Kome, S H., (2006) *Hierarchical Subject Relationships in Folksonomies*
- La Barre, K. (2006) *The Use of Faceted Analytico-Synthetic Theory as Revealed in the Practice of Website Construction and Design*, <<http://leep.lis.uiuc.edu/publish/klabarre/facetstudy.html>>.
- Morville, P., (2005) *Ambient Findability*, O'Reilly.
- Quintarelli, E., (2005) *Folksonomies: Power to the People*, Proceedings of 1' ISKO Italy-UniMIB meeting (Milano, 24 giugno 2005) <<http://www.iskoi.org/doc/folksonomies.htm>>.
- Slavic, A., (2002) *FATKS: Facet Analytical Theory in managing Knowledge Structures for humanities*, <<http://www.ucl.ac.uk/fatks>>.
- Scott, J., (2000) *Social Network Analysis*, 2nd edition, Sage Publications, London.
- Travis, W., (2006) *The strict faceted classification model*
<http://facetmap.com/pub/strict_faceted_classification.pdf>.
- Yee, K.P., Swearingen, K., Li, K., and Hearst, M., (2003) *Faceted Metadata for image searching and browsing*, Proceeding of CHI 2003 <<http://flamenco.berkeley.edu/papers/flamenco-chi03.pdf>>