

Facets are fundamental

Abe Crystal ~ abe@unc.edu ~ ASIS&T IA Summit 2006

Overview: This session reflects my own, personal frustration with the way IA is talked about. They say the best way to learn is to teach. Well, I've taught an introductory IA class for each of the past 3 summers. And I've learned that trying to explain IA processes using the standard frameworks that are out there, to students who haven't worked on large-scale information systems, can be very very frustrating. In particular, I've become disenchanted with how "standard IA" talks about organizing and representing information. I'm mad as hell, and I'm not gonna take it any more!

So I'd like to introduce the particular problems I have with existing frameworks, so you can share a bit of my frustration. And then I'll suggest that these problems all point toward making faceted classification a more fundamental part of IA processes.

My hope is to provoke you to start thinking about, and doing, IA a little bit differently. When I've presented these ideas in classes and colloquia at UNC, I've gotten lively, vigorous pushback. I hope we can continue the discussion here; I'd be happy to take questions or comments at any point.

Problem 1: "They minimize or ignore non-topical methods of organizing and presenting information."

There's been a lot of work done to integrate information architecture (which draws heavily on ILS principles of organizing and structuring

information) with user-centered design (which emerged from human factors and HCI). Despite these efforts, the existing IA frameworks are largely based on the characteristics of Web sites—not people’s information needs and behavior. These frameworks are better suited to answer questions like, “Where does this Web page fit?” than “What information does the user need, and how will she look for it?”

In my view, existing frameworks are almost perniciously hierarchy-centric and sitemap-centric. For example, Rosenfeld and Morville (2002) write that “The foundation of almost all good information architectures is a well-designed hierarchy or taxonomy” (p. 65).

These well-meaning approaches have helped bring order to the Web, but at the same time, they have placed the focus of IA on developing a topical hierarchy. However, topic or “aboutness” shouldn’t be the only focus of IA, because users look for more than topic or aboutness when they evaluate information.

In fact, an extensive body of research has investigated the criteria users rely on to assess whether a document or Web page will be relevant and pertinent to their needs. The key insight of this work is that users identify many aspects of documents as important—a whole range of criteria beyond “is this document about what I’m interested in...”

Tombros, Ruthven & Jose (2005): characteristics of Web pages, including Text, Structure, Quality, Pictures, Physical (network) Properties.

Barry (1994): characteristics of research documents, including Depth/Scope, Clarity, Novelty, Affectiveness (emotion/resonance), Availability, Source Quality, etc...

So these are two studies that illustrate the multitude of criteria that may influence users' relevance judgments and information seeking. People are making multidimensional, holistic, and highly idiosyncratic judgments.

My argument: IA frameworks and practices should incorporate these wider conceptions of relevance, move beyond a strict focus on topic and aboutness, and that faceted classification provides a way to do so.

Problem 2: "They treat facets as supplemental, rather than fundamental."

As Bates (2002) argues, faceted classification should be the foundation of Web-based information retrieval systems. Facets can be incorporated directly into information retrieval interfaces, and user studies have shown this approach be effective (Yee, Swearingen, Li, & Hearst, 2003).

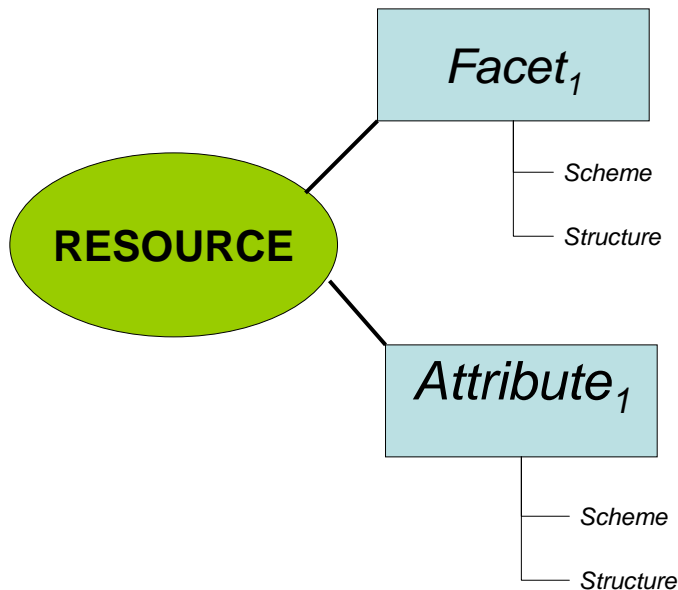
As evidenced by the numerous presentations on facet-based approaches at the 2005 IA Summit, facets are quickly becoming part of mainstream IA practice.

So what can I mean by saying that facets are being treated as supplemental, rather than fundamental?

First of all, I have the disconcerting sense that when people say “facets,” they mean “FLAMENCO.” FLAMENCO is an effective, validated interface design for faceted browsing of images. Similar designs are being used in a variety of domains, especially ecommerce. But I want to consider using facets in a broader way than this.

Second, Bates (as well as Rosenfeld and Morville) sets up an unnecessary and confusing dichotomy between hierarchical and faceted classification. As described by Yee et al (2003) and Kwasnik (1999), a facet can contain hierarchy.

FLAMENCO definition of facets: orthogonal categories; flat vs. hierarchical; single vs. multi-valued.



Schemes and Structures. Hierarchies (topic; vehicles) vs. Trees (geography; parts of a car).

Using this framework, facets can be seen as fundamental to IA, not supplemental. Rather than thinking of facets as a way to improve IA, we should think of facets as the foundation of IA. Making facets fundamental also requires clearer definitions of basic terms: what facets are, how they are chosen and defined, and how they are used.

In particular, I want to draw a distinction between facets and attributes.

Axis	Attribute	Facet
Scope	data	information
Control	syntactic	semantic
Definite-ness	tight, widely accepted	loose, controversial, problematic, culture-bound
Origin	inherent	assigned

“The conflict of intepretation is essential to the humanities.” Facets are about intepretation, so they are about conflict. They are about applying a distinct point-of-view to help make sense of large information spaces. But that viewpoint is necessarily individual, controversial, problematic.

In some cases, attributes can get messy too. But for many applications, a date is a date. A price is a price.

I hope this distinction is clear. I don't necessarily expect you to agree—these definitions have raised heated debate when I've presented them in the past.

The challenge: how to make facets fundamental.

So what?

Support faceted UI's—and innovative new forms of these UI's. Amazon diamond search—elegant visual/direct manipulation of attribute values. Much more intuitive and pleasurable to use than breaking attributes into ranges and presenting as a FLAMENCO-esque UI. Separate attributes and facets, and develop appropriate interaction designs for each.

Support non-topical IA. Earlier I argued that topic/aboutness is only one element of users' holistic relevance judgments. By emphasizing faceted organization, we open the door to non-topical facets and put these on equal footing with topic. For example, Peter Merholz has argued for the value of *genre* and suggested letting users refine by genre. A faceted model supports and advocates exactly such approaches.

Fit with database model of CMS'. But we need *better* CMS' that can support this flexible model, with both attributes and facets, and different classification schemes and structures for each facet.

Move from pages in directories to structured information objects.

Wireframe: This abstract representation (more of a page description diagram than a shows what facets are needed for a particular page (or type of page), with example information/links for each facet.

This particular rep is an example of a page or subsite targeted to a particular audience. It thus includes links to pages which are relevant to this audience. These pages might come from the AUDIENCE facet (most likely informational pages on a topic relevant to

this audience). Or, they might come from a TASK or GENRE facet. The IA selects pages which are relevant to prospective students drawing on each facet in the classification system.

Rethinking: Everybody's doing it. I began by criticizing today's IA frameworks, as embodied by the canonical Polar Bear book. I hope we'll continue to see these frameworks evolve. As IA practice matures and becomes widespread, it is important to periodically reflect on the frameworks that serve as the tacit basis of the field. The frameworks in widespread use today have helped IA's develop the field by providing a common basis for collaboration, conversation, and research. By developing more robust frameworks, we can further advance the field. I hope this presentation can serve as a first step in that work.

A further issue: “They conflate organization and representation.”

Rosenfeld and Morville, and Brinck et al both distinguish between organization and navigation systems. These distinctions simply aren't clear. In some examples, a particular set of links is called “navigation” while another set is called “organization.”

Dartmouth example. Pets.com navbar. (Think-pair-share?)

A better distinction is between the organization system created by the IA to “make order” (Levy, 1995) within the information space, and the representations that build on the organization system to make specific information objects accessible to users. A framework based on this distinction also provides an elegant way to model “hybrid” organization approaches. Rosenfeld and Morville caution against hybrid approaches, but go on to say: “the exception to these cautions against hybrid schemes exists within the surface layer of navigation... many web sites successfully combine topics and tasks within their global navigation.” In an integrated framework, these “surface layers” are not part of the organization system—they are representations of different components, combined in a way that supports users.