Tagging, User Interfaces, Taxonomies and the Semantic Web
Who I am

- Over 25 years in the business of organized information.
  - Founder, Taxonomy Strategies
  - Director, Solutions Architecture, Interwoven
  - VP, Infoware, Metacode Technologies
  - Program Manager, The Getty Trust
  - Manager, Pricewaterhouse
  - Assistant Director Technical Services, Hampshire College
  - Chief Technical Services, Paul, Weiss, Rifkind, Wharton & Garrison
- Metadata and taxonomies community leadership.
  - President, American Society for Information Science & Technology
  - Director, Dublin Core Metadata Initiative
  - Co-Founder, NKOS (Networked Knowledge Organization Systems/Services)
Recent & current projects:
http://www.taxonomystrategies.com/html/clients.htm
What I do

Organize Stuff
For us, taxonomy work includes:

- **Metadata Scheme.** Data fields for describing content so that it can be found and used.

- **Vocabularies.** Collections of terms that are used to specify some of the metadata properties.
  - Relationships between content, fields or terms (hierarchical, equivalence, & associative)
  - Some vocabularies are big & hierarchical, some are small and flat.

- **Application Profile.** Formal representation of metadata & vocabularies.
Agenda

- Content Tagging
- Tagging Interface
- Taxonomies
- Semantic Web
Tagging overview

- Tagging is better than the words that happen to occur in a piece of content.
- All tagging is useful
  - End user tagging
  - Tagging by librarians
  - Automated tagging by OS and algorithms
- Content should be tagged throughout its lifecycle, each time the content is handled and used so that it accrues value or its significance is diminished.
MS Office: File → Properties

How many people fill this in?
How many people click on this?
What is social tagging?

- End user tagging
- Easy, intuitive tagging interfaces
- Almost instantaneous feedback
  - Enables people to tag & re-tag content
  - … in response to seeing their tags in context with other tags.
- Emergent categories
  - Resembles open card sort process in which patterns emerge
  - … rather than validating categories using closed card sorts.
Some social tagging innovators

- del.icio.us founder (2003)
  - Joshua Schachter
  - Caterina Fake
  - Stewart Butterfield
  - Mark Zuckerberg
- You Tube founders (2005)
  - Steve Chen
  - Chad Hurley

- del.icio.us and flickr were acquired by Yahoo in 2005; YouTube was acquired by Google in 2006.
- As of Nov 2007 flickr had 2 billion photos posted.
Agenda

- Content Tagging
- Tagging Interface
- Taxonomies
- Semantic Web
Requirements for a tagging interface

- Automated form fill-in: Automatically fills in known data.
- Tagging precedents: See tags already assigned by others.
- Controlled vocabularies: Pull-down list.
- Multi-valued tags
- Geo-tagging
- Group tagging
  - Batch operations
  - Inherit properties from a group
- Clean-up tag tools
  - Alpha list
  - Global changes
- Batch editing
- Privacy setting: Share/Don’t share.
- Almost immediate feedback
  - Most popular
  - Tag cloud
  - Alerts
Form fill-in:
Automatically fills in known data

More detail about Carsten Höller Test Site, Tate Modern, Dec 2006

![Image](image_url)

**Taken on**
December 1, 2006 at 8:32am PST

**Posted to Flickr**
December 7, 2006 at 10:47am PST

<table>
<thead>
<tr>
<th>Camera:</th>
<th>Canon PowerShot SD110</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure:</td>
<td>0.125 sec (1/8)</td>
</tr>
<tr>
<td>Aperture:</td>
<td>f/2.8</td>
</tr>
<tr>
<td>Focal Length:</td>
<td>5.4 mm</td>
</tr>
<tr>
<td>Exposure Bias:</td>
<td>0/3 EV</td>
</tr>
<tr>
<td>Flash:</td>
<td>Flash did not fire</td>
</tr>
</tbody>
</table>

| X-Resolution: | 180 dpi |
| Y-Resolution: | 180 dpi |
| Date and Time: | 2006:12:01 08:32:33 |
| YCbCr Positioning: | Centered |
| Date and Time (Original): | 2006:12:01 08:32:33 |
| Date and Time (Digitized): | 2006:12:01 08:32:33 |
| Compressed Bits per Pixel: | 5 bits |
| Shutter Speed: | 96/32 |
| Maximum Lens Aperture: | 95/32 |
| Metering Mode: | Pattern |
| Color Space: | sRGB |
| Focal Plane X: | 9480.154 dpi |
| Focal Plane Y: | 9480.154 dpi |
| Sensing Method: | One-chip colour area sensor |

What is EXIF data?

Almost all new digital cameras save JPEG (jpg) files with EXIF (Exchangeable Image File) data. Camera settings and scene information are recorded by the camera into the image file. Examples of stored information are shutter speed, date and time, focal length, exposure compensation, metering pattern and if a flash was used.

Source: [DigitalHelper](http://www.digitalhelper.com)
Tagging precedents:
See tags assigned by others
Controlled vocabularies: Pull-down lists
Multi-valued tags

Separate each tag with a space: cameraphone urban moblog. Or to join 2 or more words together in one tag, use double quotes: "eric is the best".
Geo-tagging
Group tagging: Batch operations

Describe your photos
Or, open in Create for more fine-grained control.

Batch operations

Add Tags

"Clarin Alley" "San Francisco" Murals
ADD

Add to a Set

Choose a Set

Choose a Set
Alemany Farmers Market
Tate Slides

Titles, descriptions, tags

Title:
Mariachis

Description:

Title:
Virgin of Guadalupe

Description:

Title:
Guernica

Description:
Group tagging: Tag a folder

- Drag & drop content items into folder (or group)
- Then, content items inherit properties of folder

Example: iTunes music inherits properties of album, such as artist, genre, etc.
Clean-up tools: Alpha list

<table>
<thead>
<tr>
<th>Tag</th>
<th>Entered as</th>
<th>On</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>alemanymarket</td>
<td>Alemany Market</td>
<td>6 photos</td>
<td>Edit</td>
<td>Delete</td>
</tr>
<tr>
<td>farmersmarket</td>
<td>Farmers Market</td>
<td>7 photos</td>
<td>Edit</td>
<td>Delete</td>
</tr>
<tr>
<td>farmersmarket</td>
<td>Farmers Market</td>
<td>5 photos</td>
<td>Edit</td>
<td>Delete</td>
</tr>
<tr>
<td>flowers</td>
<td>Flowers</td>
<td>1 photo</td>
<td>Edit</td>
<td>Delete</td>
</tr>
<tr>
<td>murals</td>
<td>Murals</td>
<td>7 photos</td>
<td>Edit</td>
<td>Delete</td>
</tr>
<tr>
<td>sanfrancisco</td>
<td>San Francisco</td>
<td>7 photos</td>
<td>Edit</td>
<td>Delete</td>
</tr>
</tbody>
</table>
Clean-up tools: Global changes

- ID collection of related content items by pattern or context
- Then, apply same attributes to all content items

**Tags**

- rename tags
- delete tags
- bundle tags

**Tags: rename tags**

This is where you can rename tags that you have used before. Renaming tags will change every bookmark that is tagged with the old tag to the new tag. You can also delete tags.

- old tag: [Choose a tag to rename]
  - one or more new tags
  - replace tags
Batch editing
Privacy settings: Share/Don’t share

url: http://www.acm.org/class/
description: ACM Computing Classification Systems
notes:
tags: vocabulary computing classification ACM

save

originally posted on 2006-08-27. delete this post.
Almost instantaneous feedback: Interactive rewards

- Almost instantaneous exposure of tags in simple user interfaces on the web provides positive reinforcement for user tagging that simply did not exist before.

- For example,
  - Most popular
  - Tag clouds
  - Alerts
Most popular

Hot tags

In the last 24 hours
itunes7, itv, lunch20, addresscontinenteurope, roidweek06, coverflow, showtime, itunes, mobilife, contextwatcher, shuffle, primary, girona, puebla, bestival, fz7, cellmcc310, nano,theworldthroughmyeyes, viewfinder

Over the last week
utatathursdaywalk21, artparade, podcamp, towersoflight, torontointernationalfilmfestival, asem, zombiepubcrawl, zombiepubcrawl2006, openmonumentendag, solanostroll, bestival2006, virginfestival, tributeinlight, bestival, dconstruct06, deitch, dconstruct, touchandgo, leuchars, islademaipo

* Another example is most emailed from, e.g., the NY Times.*
Alerts

- New (content selected by date)
- Subscriptions (content selected by tags)
- Interest (content selected by other people)
- Individual (content selected for you by other people)

People: subscriptions

This is where you set up your tag subscriptions, which allow you to keep track of all new bookmarks saved with tags that interest you. View your subscriptions page at del.icio.us/subscriptions/jousch.

subscribe to a tag

tag
only from this user (leave blank for all)

subscribe

label subscription
Summary

- There are lessons to be learned from web tagging about how to get good metadata.
- Tagging must be simple, and it must be almost instantaneously easier to find relevant stuff.
Agenda

- Content Tagging
- Tagging Interface
- Taxonomies
- Semantic Web
Taxonomy trends on the Web

- Lesson from shopping websites – a few small taxonomies, rather than one large one.
  - Easier to tag
  - Easier to build services
  - Easier to maintain
Hi all,

I am forwarding a question from my department head, in hopes that you may have leads on this topic.

Does anyone have experience with material that is being produced on campuses (teaching materials, lab experiments, etc.) and uploaded to youtube. are any of you aware of staff that is cataloging this material? or any discussion around this in the larger community? seems valuable to capture and expose.....

Thanks,

Louise Ratliff

UCLA Library Cataloging & Metadata Center
You Tube metadata

- Channel
- Title
- Audience
- Creator
- Description
- Date
- Category
- Subject
- Identifier
- Format
- Relations
- Rating
- Comments

Dublin Core
Other attributes
You Tube taxonomies

Channel
- All
- Comedians
- Directors
- Gurus
- Musicians
- Non-Profit
- Partners
- Sponsors
- YouChoose 08

Audience
- Autos & Vehicles
- Comedy
- Education
- Entertainment
- Film & Animation
- Howto & Style
- Music
- News & Politics
- Nonprofits & Activism
- People & Blogs
- Pets & Animals
- Science & Technology
- Sports
- Travel & Events

Category
- Autos & Vehicles
- Comedy
- Education
- Entertainment
- Film & Animation
- Howto & Style
- Music
- News & Politics
- Nonprofits & Activism
- People & Blogs
- Pets & Animals
- Science & Technology
- Sports
- Travel & Events

Rating
- Featured
- Most Subscribed
- Most Viewed
- Most Discussions
- Most Members
- Most Videos
- Most Active
- Most Discussed
- Most Recent
- Most Members
- Most Responded
- Most Viewed
- Previously Popular
- Top Favorites
- Top Rated

Audience & Category use the same controlled vocabulary
You Tube services:
From just a little bit of tagging

- Populate all Video, Channel & Community pages.
- Automatic RSS subscriber based on tags for any page.
- Identify more videos from the same creator.
- Identify related videos.
- Create, publish & monetize (place ads) your playlists & favorites.
Agenda

- Content Tagging
- Tagging Interface
- Taxonomies
- Semantic Web
What is the semantic web

- Making content web-accessible in a format that can be read and used by automated tools, so that people and machines can find, share and integrate information more easily.

- Some current examples, especially if they use semantics as the basis for that integration.
  - **Dynamic Web Pages** – Content changes in response to different contexts or conditions.
  - **Personalization** – Tailoring to a user based on personal details or characteristics they provide.
  - **Mashups** – Combining data from more than one source into an integrated application.
Oracle events
Technology: Siderean
http://events.oracle.com/

- Traditional search
- Filter on Location “Germany” and Date “This Month”
- Time line search widget
- More filters based on this result
- Results shown on Google maps UI
- Subscribe to RSS feed based on the criteria set on this page

Register, view details, or email info
Oracle top-level taxonomy architecture

“Is a” groups of Products
Questions?

Joseph A. Busch
415-377-7912 jbusch@taxonomystrategies.com