Tagging, User Interfaces, Taxonomies and the Semantic Web

Copyright 2008 Taxonomy Strategies LLC. All rights reserved.

Who I am

http://www.taxonomystrategies.com/html/josephbusch.htm

- 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:

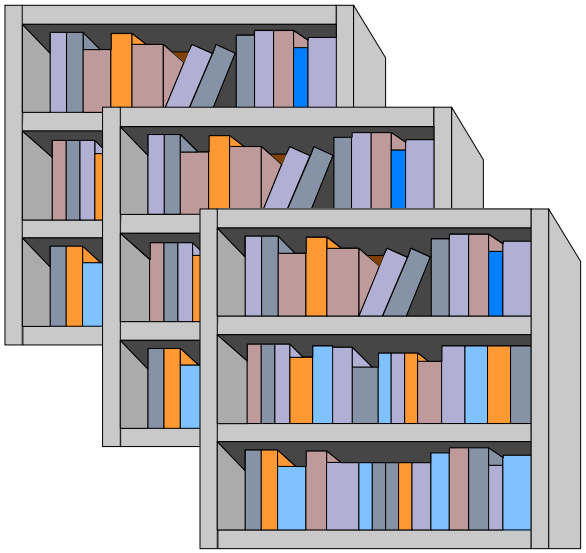
http://www.taxonomystrategies.com/html/clients.htm

Government

Commercial

Not-for-Profit

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.

Element	Data Type	Length	Req. / Repeat	Source	Purpose
Creator	String	Variable	?	System supplied	Content ownership.
Organization	String	Variable	1	User supplied	
Dates	Date	Fixed	1	System supplied	Publish, feature, & review content.
Description Metadata					
Content Type	String	Variable	1	Content Type CV	Localization & deriving new items.
Page Name	String	Variable	*	Validate by lookup	Reference to related resource.
Version	String	Variable	*	Validate by lookup	
Title	String	Variable	1	System supplied	Text search & results display
Description	String	Variable	1	User supplied	
Use Metadata					
Segment	String	Variable	*	Segment CV	Target, personalize content.
Product	String	Variable	*	SKU CV	
Locale	String	Fixed	1	Culture Code CV	Language & country localization.
Subject Metadata					
	String	Variable	*	Task CV	Target, personalize content
	String	Variable	*	RML CV	
Keywords	String	Variable	*	Editor supplied	Text search & results display

Legend: ? – 1 or more * - 0 or more

	A	B	C	D	E	F	G	H
1	Content Types						Synonyms	Notes
2								
3		Calendar						Role-specific overview of major milestones.
4		Checklist						Role-specific laundry list of to-do items with links
5		Executive Brief						Outlines management issues for a given topic an
6		How-to						Enumerates steps for accomplishing a specific ta
7		Insight Editorial						Provides thought leadership; generally issue orier
8		Marketplace Page						Listing of ISVs, partners, or related products.
9		Multimedia Page						Introductory text and links to videos, demos or fla
10		Product Capability Study						Details a single area of functionality as it relates
11		Product Comparison						Compares functionality between different editions
12		Product Feature Definition						Brief description and business benefit of an indivi
13		Product Overview						High-level survey of a Microsoft product.
14		Product UI						Snapshot of a product UI with labels for individual
15		Quiz						Multiple choice questions with the ability to make
16		Solution Overview						High-level narrative about how a customer solves
17		Special Offer						Lead generation page focused on product, campa
18		Step-by-Step Guide						Explains the high-level steps in a business proce
19		System Requirements						Configuration-specific hardware and software requ
20		Talking Points						Key issues BDMs should know about when enga
21		Theme Page						Aggregated links to content about a single topic.
22		Tip						Short step to follow that yields immediate ease o
23								
24								
25								
26								
27								
28								
29								

Content Types / Audience / Culture Codes / Tasks / RML / SKU / Sheet1

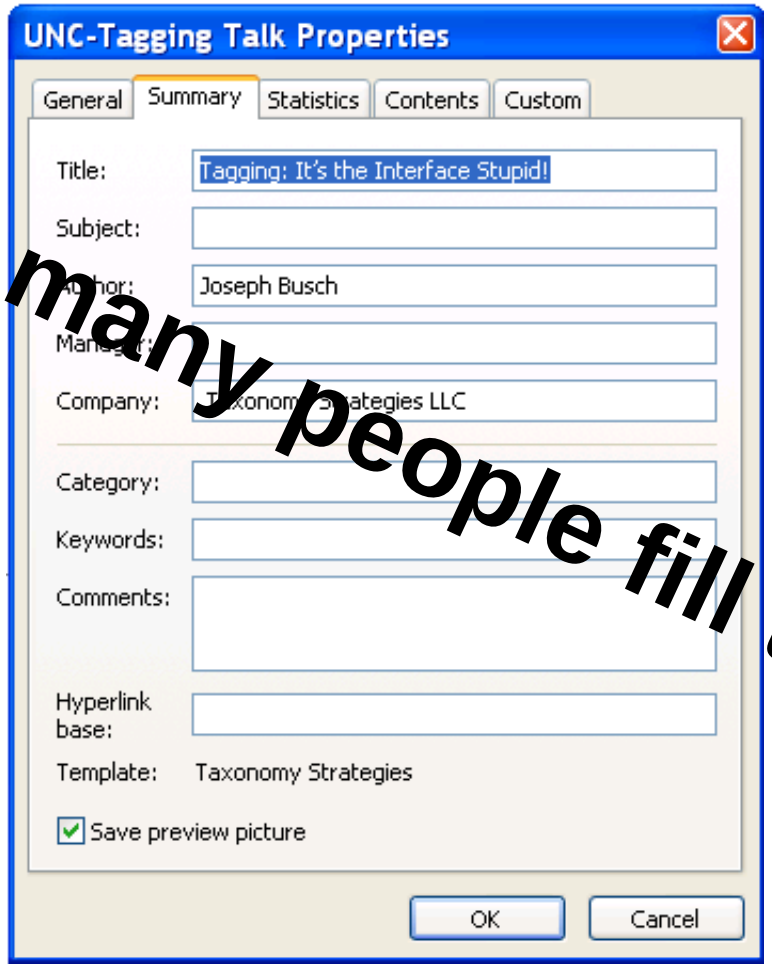
Agenda

- 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?

Organize

How many people click on this?

Activity On Your Photos | Comments You've Made | In Your Groups | Photos from your friends

You Your Photos | Organize | Upload | Your Account | Do More with Your Photos

Explore Places | Last 7 Days | This Month | Popular Tags | Creative Commons | Search

Help Community Guidelines | The Help Forum | FAQ | Tools | Sitemap | Help by Email

Send to a friend

Save to del.icio.us

Flickr Blog | About Flickr | Terms of Service | Your Privacy | Copyright/IP Policy | Report Abuse

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
- flickr founders (2004)
  - Caterina Fake
  - Stewart Butterfield
- facebook founder (2004)
  - Mark Zuckerberg
- You Tube founders (2005)
  - Steve Chen
  - Chad Hurley

- del.icio.us and flickr were acquired by Yahoo 2005; YouTube was acquired by 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

Taken on
December 1, 2006 at 8.32am PST

Posted to Flickr
December 7, 2006 at 10.47am PST

Edit the photo dates

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: Digicamhelp.

Camera:	**Canon PowerShot SD110**
Exposure:	**0.125 sec (1/8)**
Aperture:	**f/2.8**
Focal Length:	**5.4 mm**
Exposure Bias:	0/3 EV
Flash:	Flash did not fire
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-Resolution:	9846.154 dpi
Focal Plane Y-Resolution:	9846.154 dpi
Sensing Method:	One-chip colour area sensor

Tagging precedents:
See tags assigned by others

del.icio.us

TagCloud - Home
http://www.tagcloud.com/
this url has been saved by 2117 people.

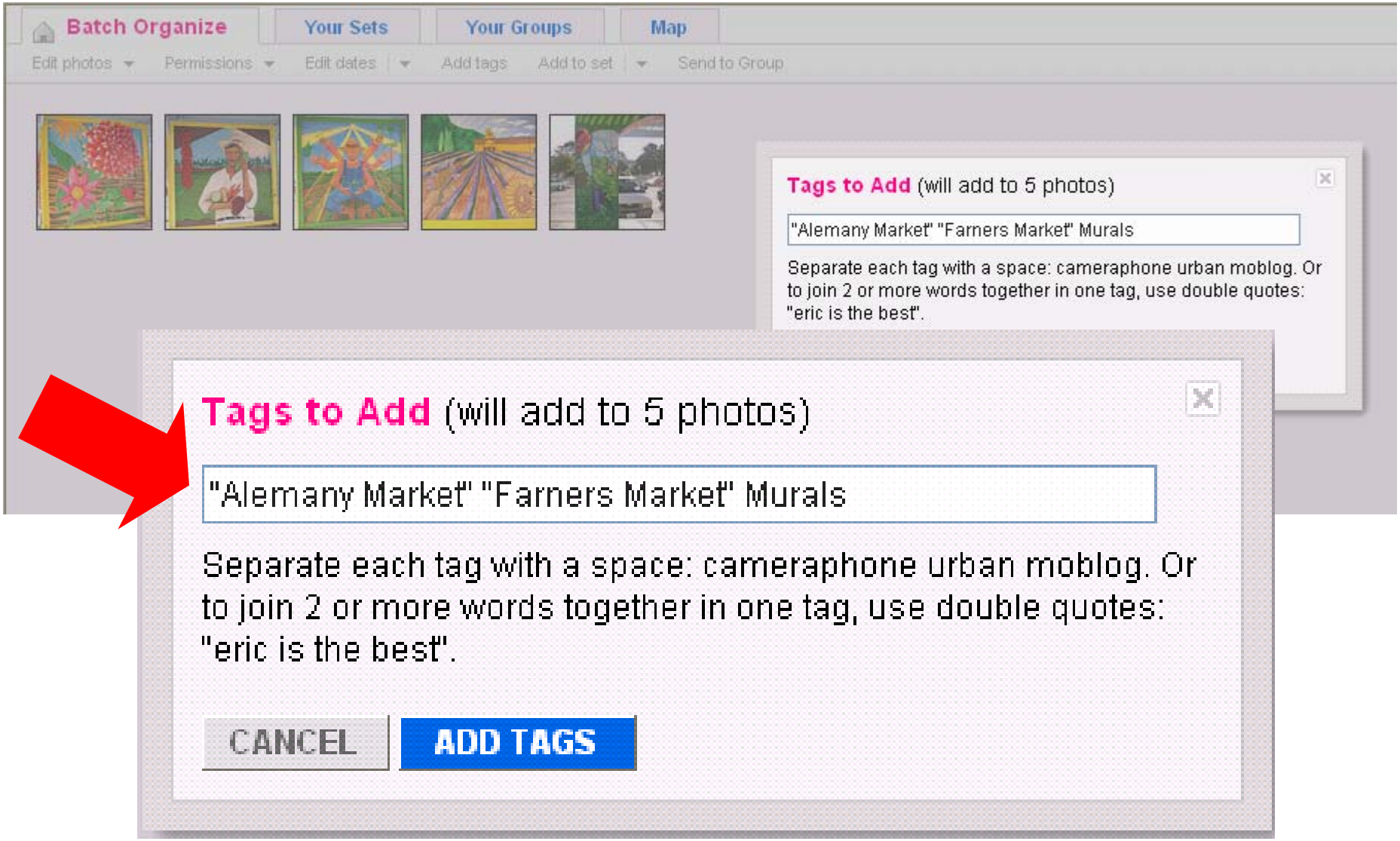
title TagCloud - Home
url http://www.tagcloud.com/
notes
tags tagging folksonomy
save cancel do not share full-screen edit

common tags cloud | list

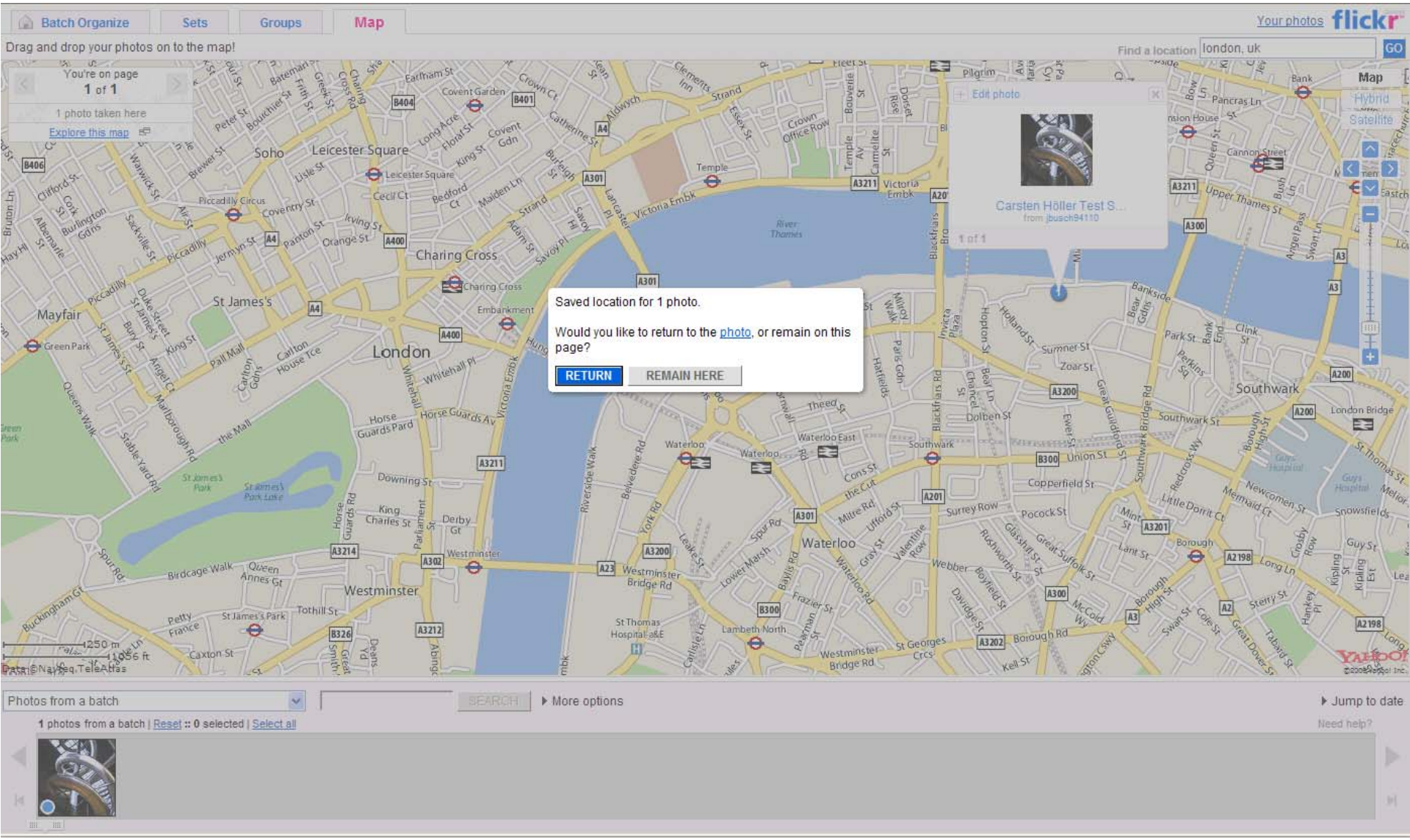
aggregator ajax blog blogging blogs cloud del.icio.us
design folksonomy internet news rss search
social software tag tagcloud tagging tags tool
tools visualization web web2.0 webdesign

Controlled vocabularies:
Pull-down lists

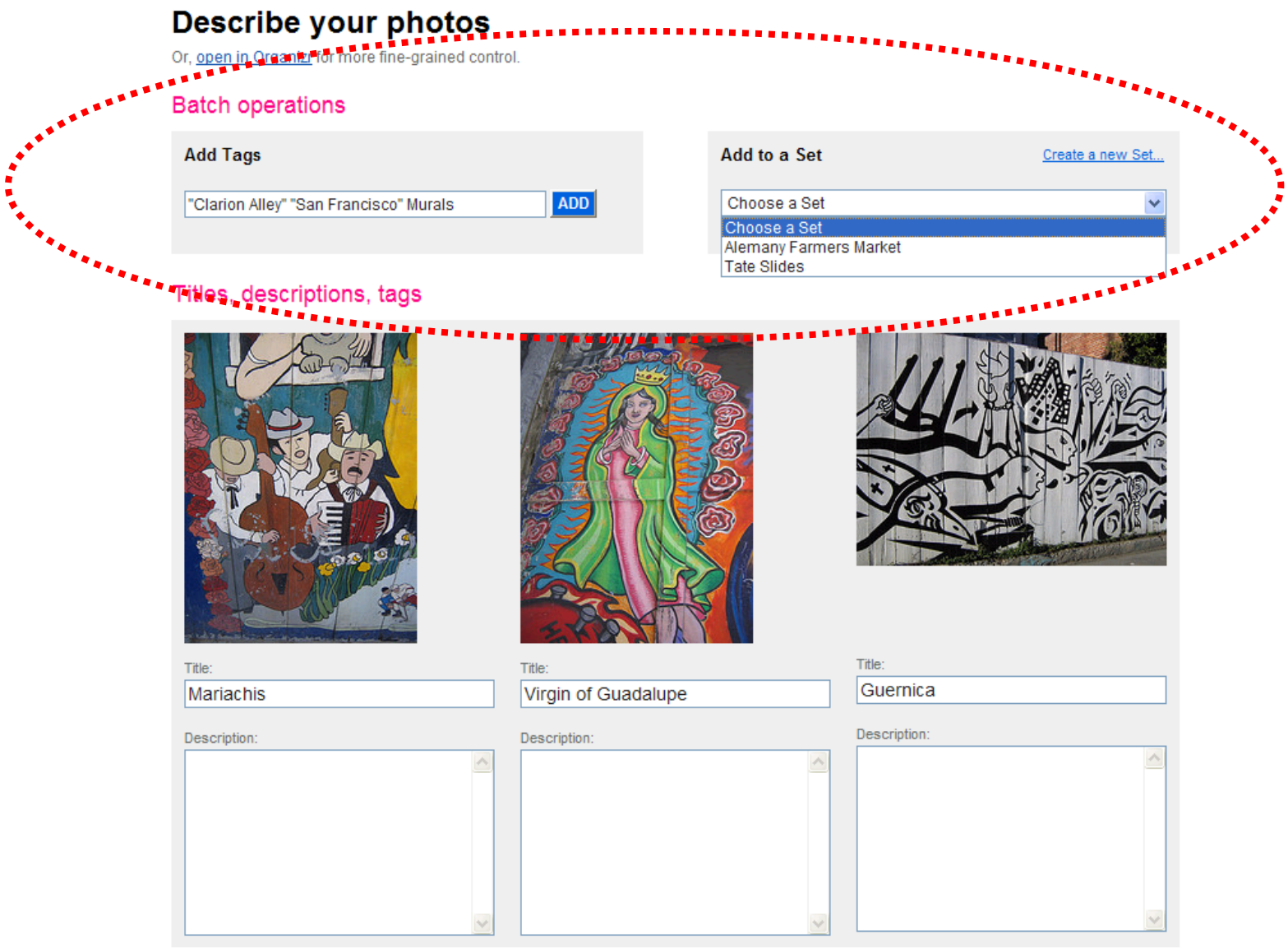
Multi-valued tags



Geo-tagging


Batch Organize
Sets
Groups
Map
Your photos flickr
Drag and drop your photos on to the map!
Find a location london, uk
GO
You're on page 1 of 1
1 photo taken here
Explore this map
Edit photo
Carsten Höller Test S...
from jbusch94110
1 of 1
Map
Hybrid
Satellite
Saved location for 1 photo.
Would you like to return to the photo, or remain on this page?
RETURN
REMAIN HERE
Soho
Leicester Square
Piccadilly Circus
Charing Cross
Mayfair
St James's
London
Westminster
Waterloo
Southwark
Photos from a batch
SEARCH
More options
Jump to date
1 photos from a batch | Reset :: 0 selected | Select all
Need help?

Group tagging: Batch operations

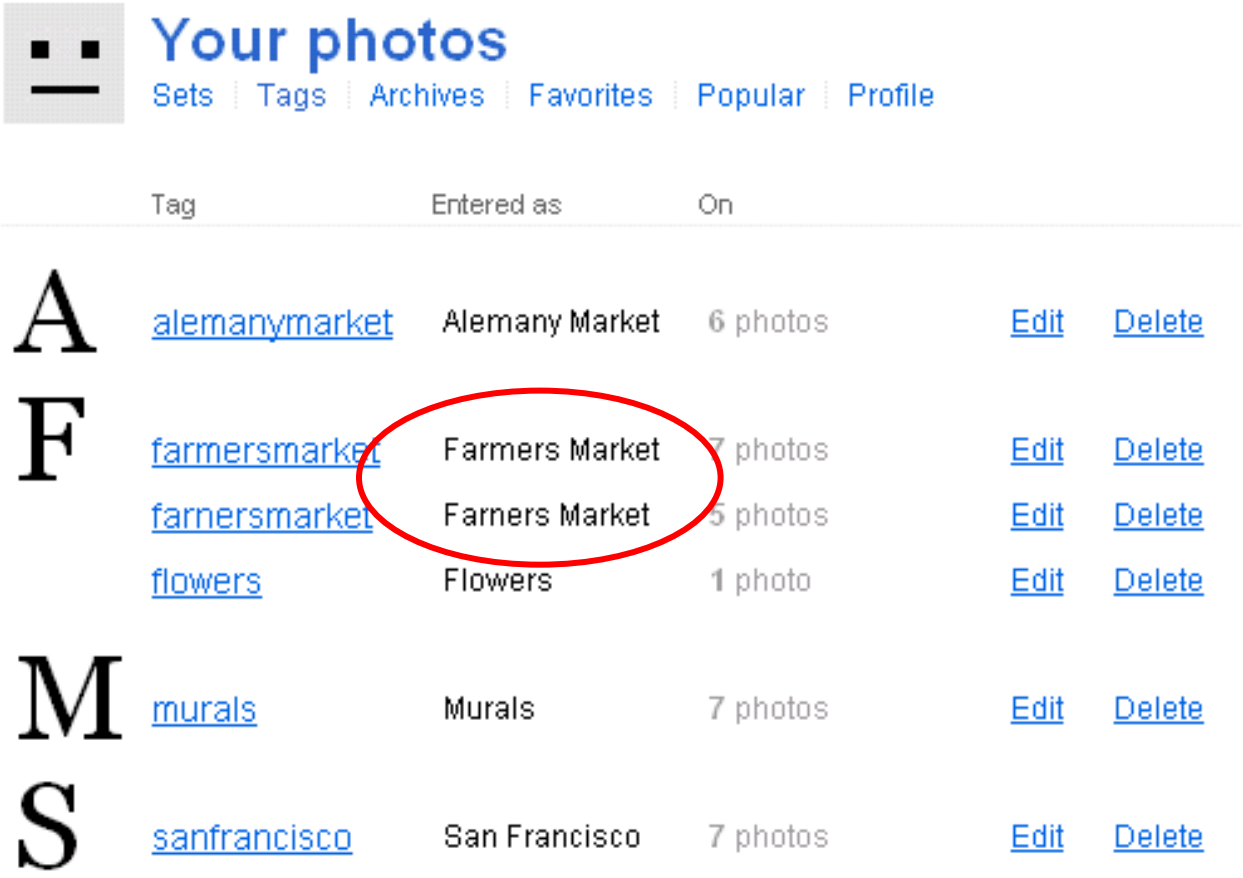


Group tagging: Tag a folder

- 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



Clean-up tools: Global changes

del.icio.us

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

Tags: rename tags

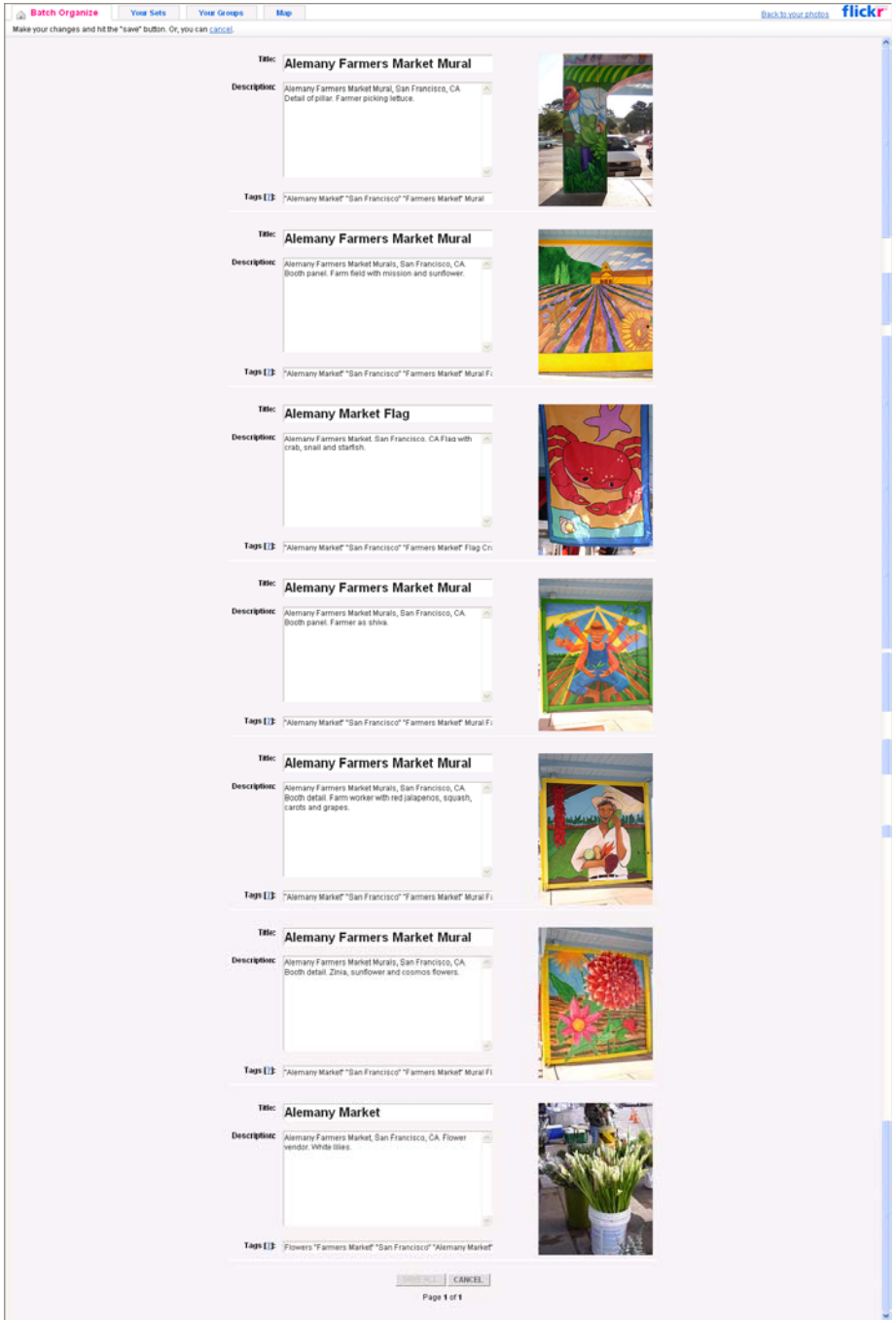
This is where you can rename tags that you have used before. Renaming tags will change every bookmark that is tagged with the old tag to the new tag. You can also delete tags.

old tag -- choose a tag to rename --

one or more new tags

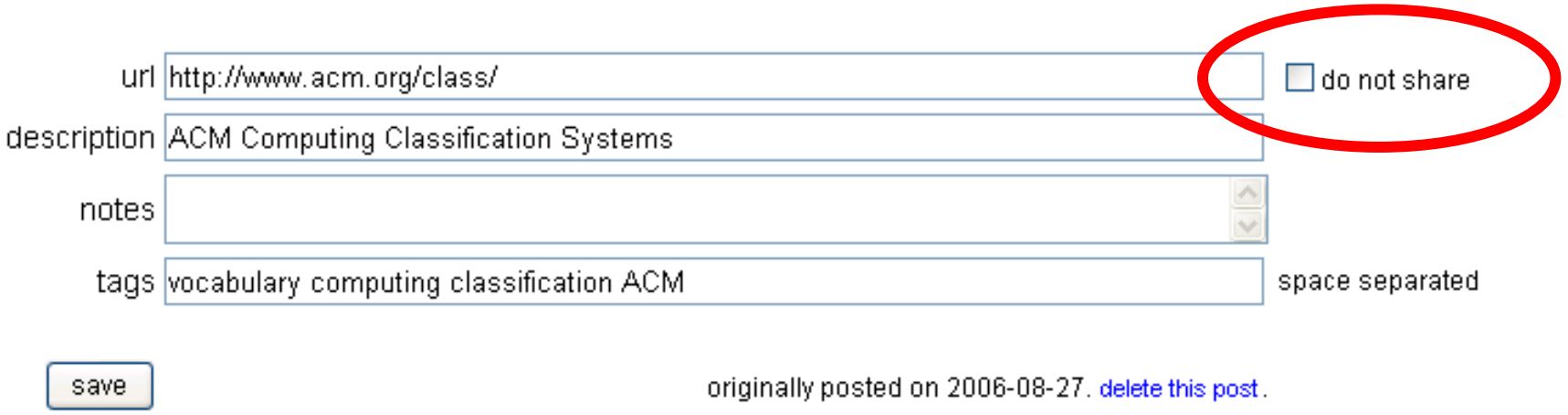
replace tags

Batch editing



Privacy settings:
Share/Don't share

del.icio.us



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.**

Tag cloud

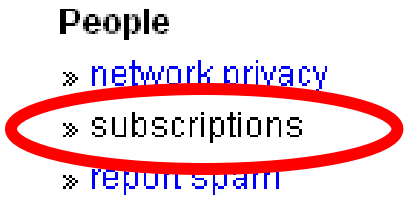
All time most popular tags

06 africa amsterdam animal animals april architecture art august australia baby barcelona beach berlin birthday black blackandwhite blue boston bw california cameraphone camping canada canon car cat cats chicago china christmas church city clouds color concert day dc dog england europe family festival film florida flower flowers food france friends fun garden geotagged germany girl graffiti green halloween hawaii hiking holiday home honeymoon hongkong house india ireland island italy japan july june kids lake landscape light live london losangeles macro march may me mexico mountain mountains museum music nature new newyork newyorkcity newzealand night nikon nyc ocean paris park party people portrait red river roadtrip rock rome san sanfrancisco school scotland sea seattle show sky snow spain spring street summer sun sunset sydney taiwan texas thailand tokyo toronto travel tree trees trip uk urban usa vacation vancouver washington water wedding white winter yellow york zoo

Alerts

del.icio.us

- 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 keep track of all new bookmarks saved with tags that interest you. View your subscriptions page at del.icio.us/subscriptions/jbusch.

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.

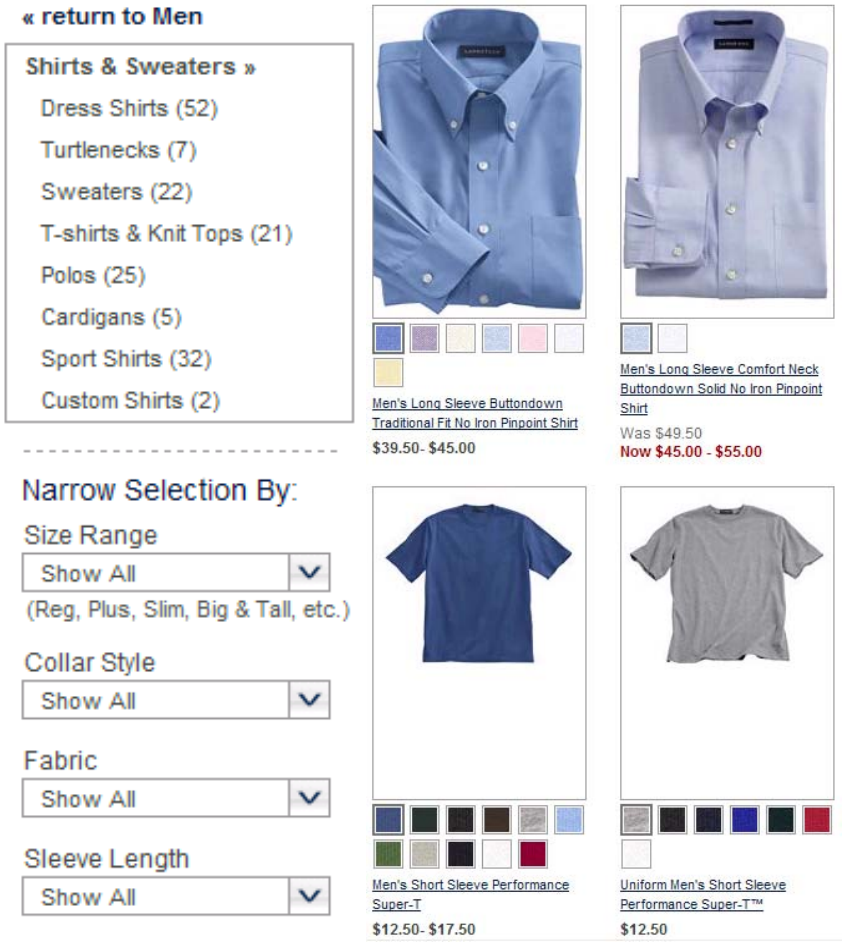
Agenda

- Content Tagging
- Tagging Interface
- Taxonomies
- Semantic Web

Taxonomy trends on the Web

LANDS' END

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



Date: Wed, 6 Feb 2008 17:23:30 -0800
From: lratliff@library.ucla.edu
To: metadatalibrarians@lists.monarchos.com
Subject: [Metadatalibrarians] Cataloging of videos on YouTube

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

Metadatalibrarians mailing list
Metadatalibrarians@lists.monarchos.com
http://lists.monarchos.com/listinfo.cgi/metadatalibrarians-monarchos.com

You Tube metadata

Dublin Core

Other attributes

Channel

Audience

Title

Creator

Description

Date

Category

Subject

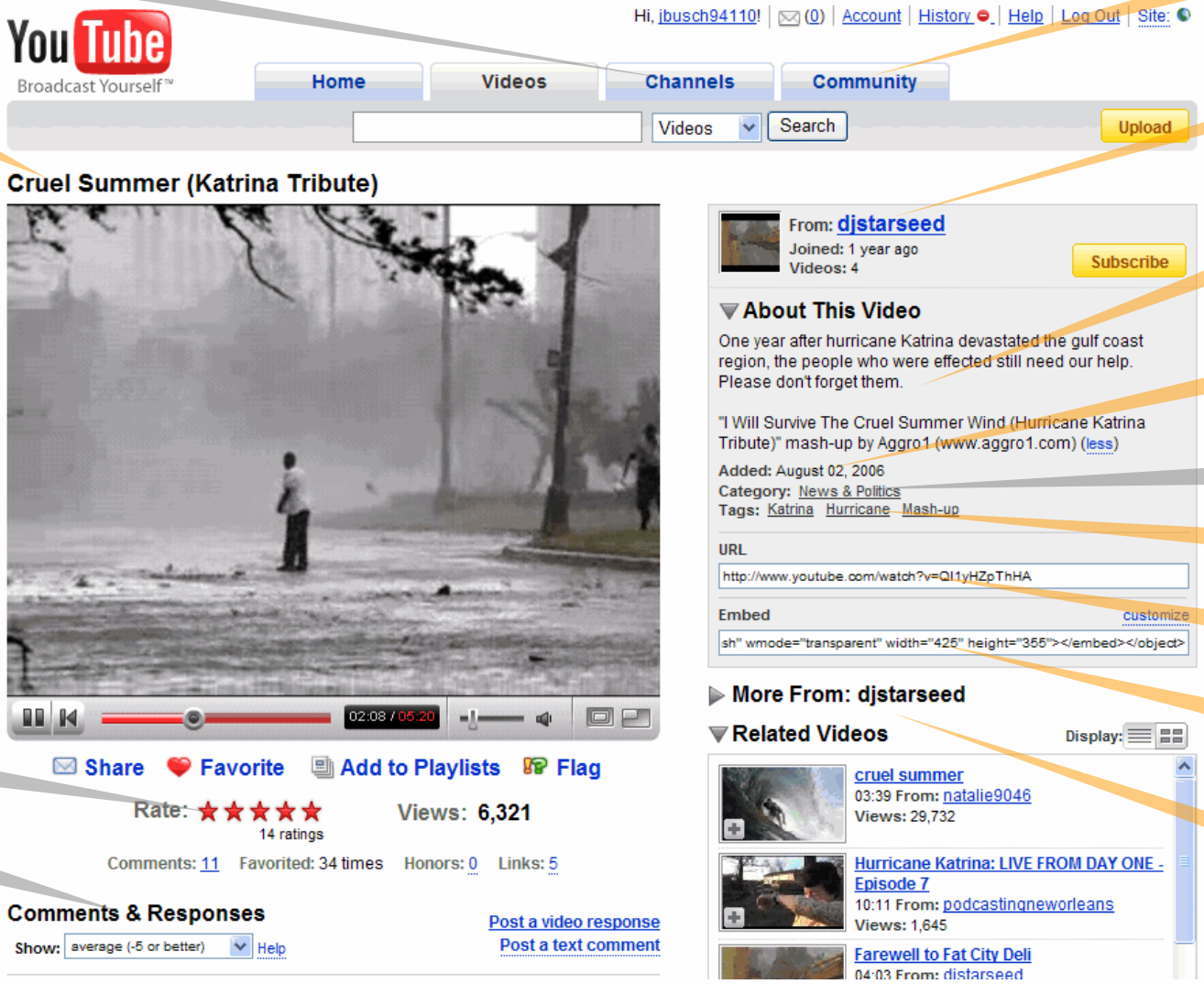
Identifier

Format

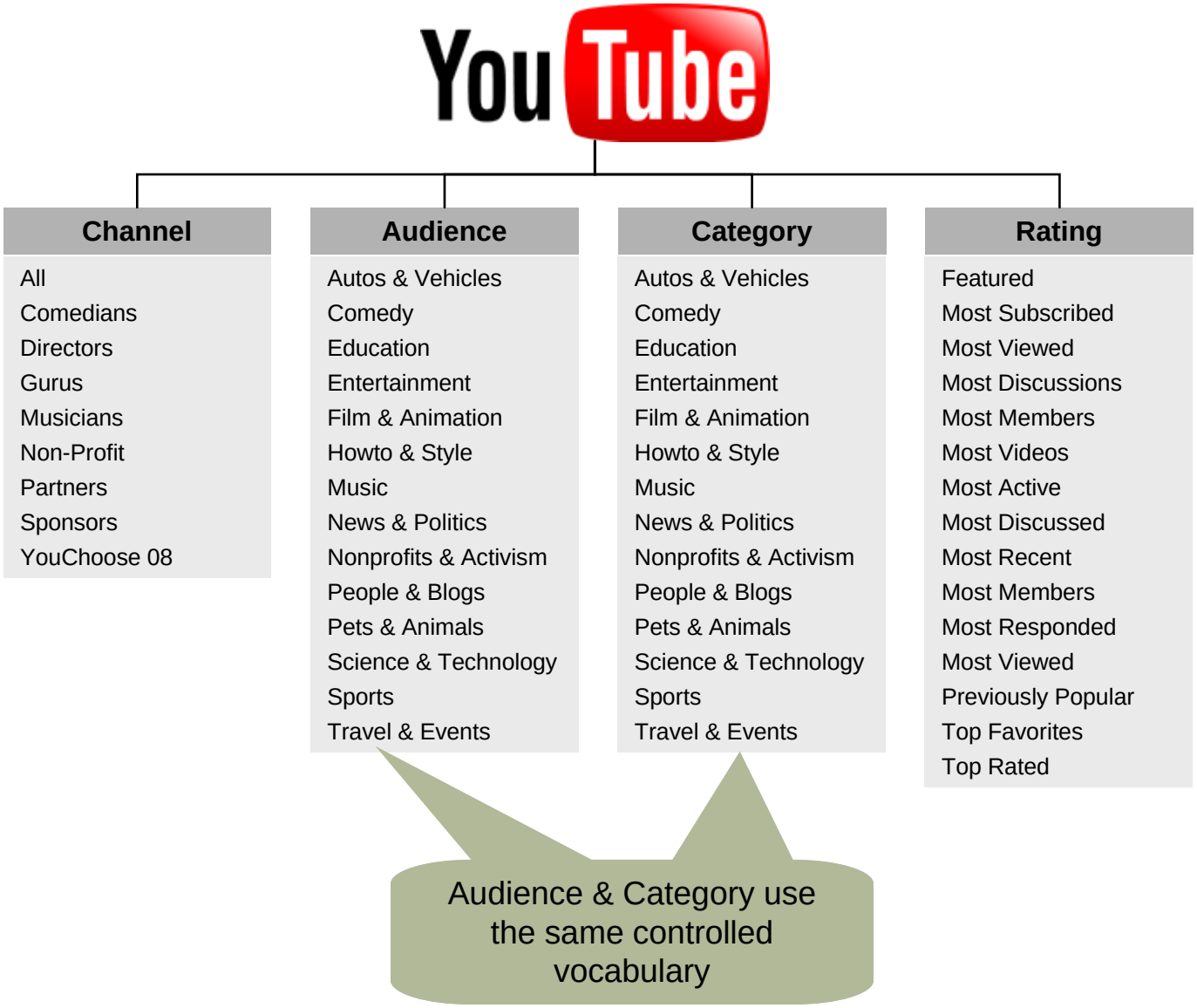
Relations

Rating

Comments



You Tube taxonomies



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

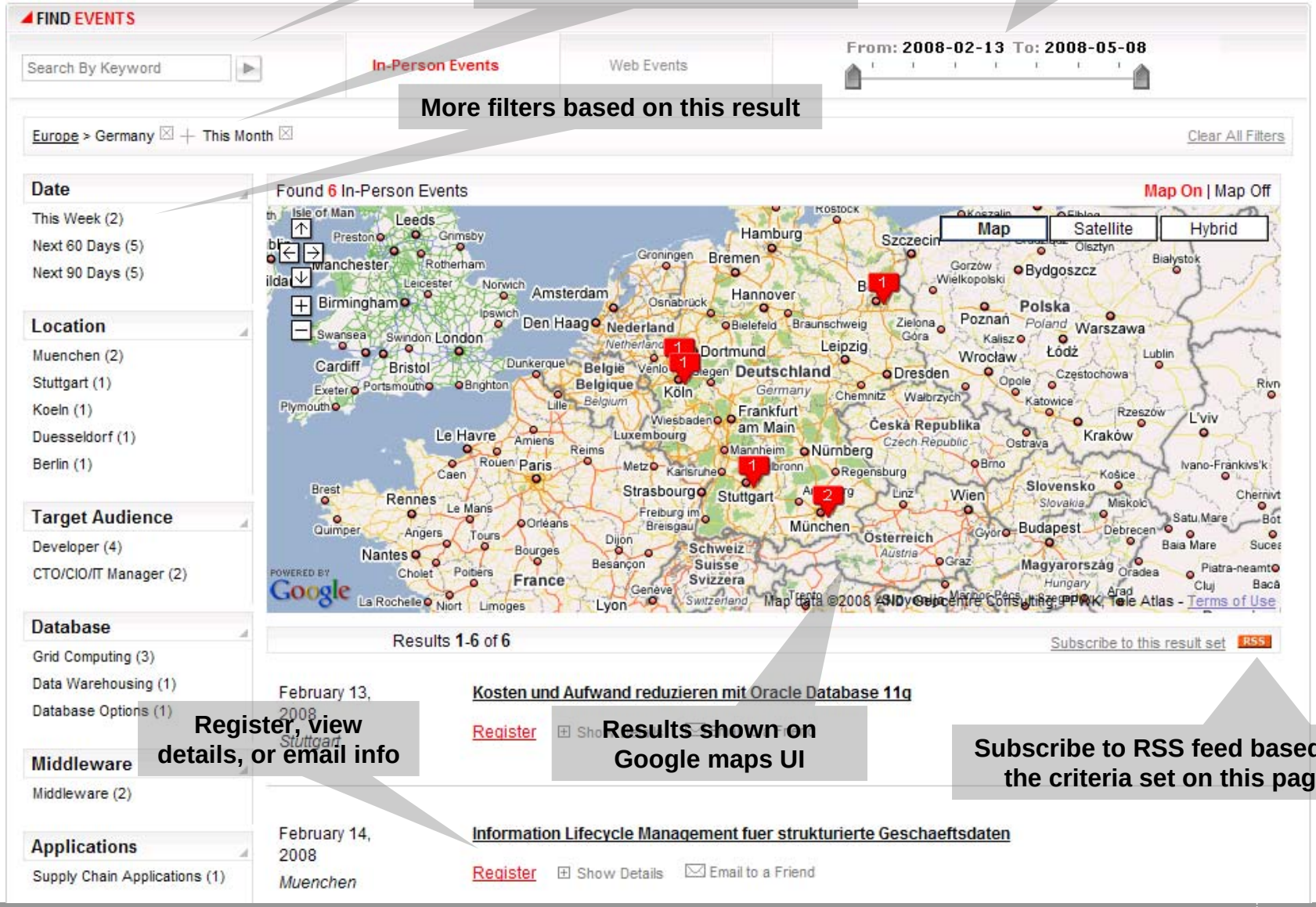
http://events.oracle.com/

Traditional search

Filter on Location "Germany" and Date "This Month"

Time line search widget

More filters based on this result



Register, view details, or email info

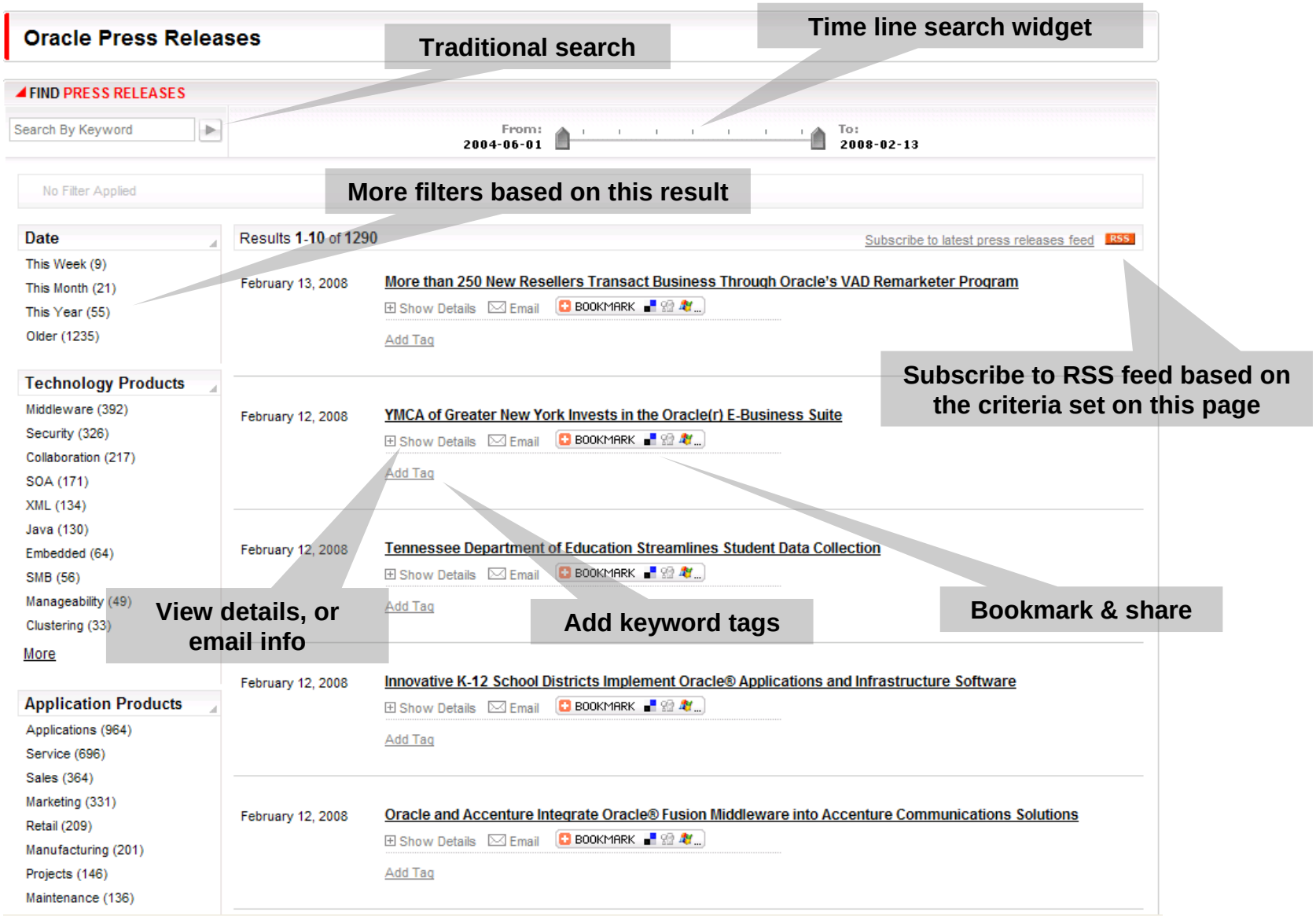
Results shown on Google maps UI

Subscribe to RSS feed based on the criteria set on this page

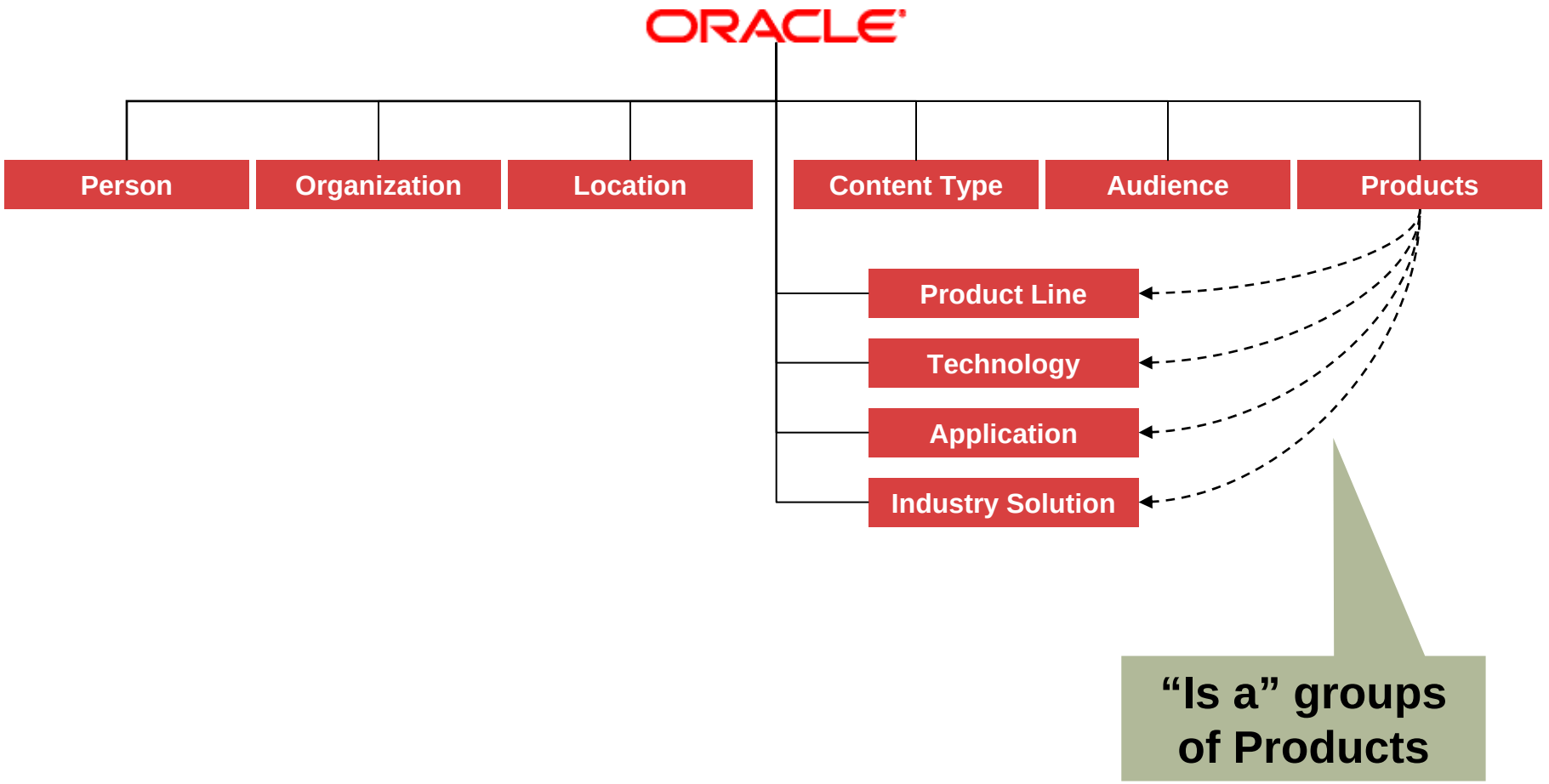
Oracle press releases

Technology: Siderean

http://pressroom.oracle.com/



Oracle top-level taxonomy architecture



Questions?

Joseph A. Busch

415-377-7912 jbusch@taxonomystrategies.com

Copyright 2008 Taxonomy Strategies LLC. All rights reserved.