Beyond Buckets:
An Introduction to
Drupal Taxonomy

Sarah Prown
prown@reedsparrow.com
What is taxonomy?

- Classification
- Organization
- Description
- Access Points
- Intellectual Control
How is taxonomy used?

- Scientific classification
- Libraries (subject headings)
- Your sock drawer?
- Taxonomy is used any place you want to exert intellectual control over a diverse collection of information (or objects, or ideas, or...)

Taxonomy in Libraries

- Libraries use many ways of organizing their collections
  - Author
  - Format
    - DVDs are separate from magazines
  - Subject
- Now, imagine a library where the materials were organized by color?
- Or size?
Who remembers this?
Or this?
Everyone loves this...
Thanks for the library lesson, but...

- What does this have to do with my drupal site?
- Let's take it back a step, and talk about...
What is a website?

- It's a collection of information/data/etc.
- It's a tool.
- It's an experience.
- Its success or failure is determined by one user's ability to find what they need at the time they need it...
Beyond Buckets

- Think about your favorite websites...
- What works?
  - Transparency
  - Discovery
  - Serendipity

➡ Users invited in to explore, make connections.
Buckets Buckets Buckets

- Content Management Systems encourage bucket-thinking
  - Content Types
  - Architecture
  - Menus
  - Hierarchies
- These are all essential things, by the way...
- A strong internal organization and structure is always the first step
Another dimension...

- Taxonomy allows you to rethink/reorganize content in new ways.
- New connections
- New pathways to content (no orphans!)
- Serendipity and Discovery
- Flattening of hierarchy
- Tag clouds
- User tagging (free-tagging, community)
Taxonomy in Action: Drupal

- Modules Search
- Exposed Filters
- Select Any and browse all.
- Faceted browsing
The YDC2 Website

- Challenges:
  - A relatively new department at Yale
  - Marketing (who we are/what we do)
  - Education
  - Communications

- Taxonomy allowed us to address these challenges
About YDC2

YDC2 - the Yale Digital Collections Center, formerly the Office of Digital Assets and Infrastructure, has grown out of an initiative of the Provost to develop a digital content strategy for Yale University.

YDC2 has three components that subsume most of the activities previously supported by ODAI and adds greater focus on the development of digital support strategies for the Collections Study Center on West Campus.

Mission:

The Yale Digital Collections Center (YDC2) at West Campus makes information about Yale's preeminent collections digitally accessible to the world. By supporting Yale in the identification, development, and implementation of shared digital tools and methods, YDC2 facilitates new inquiry through cultural heritage and natural science collections and further advances conservation practice through data-driven methods. Leveraging the Collections’ activity and scientific resources of West Campus, YDC2 enables faculty and students within the global community to use digital collections and information about the collections in innovative ways. YDC2 provides strategic planning and project coordination for the museums and the Arts Area Schools to build Yale’s capacity to build Yale’s capacity to 1) capture, synthesize and analyze collections and conservation data, 2) manage digital content and provide access over time, 3) use digital technologies to improve conservation of and research into collected objects.

Components: Create, Manage, Use

1. Imaging Lab: A shared digital and scientific imaging laboratory on West Campus (formerly known as Digital Core) to extend capacity in digital photography and introduce innovative methods of scientific imaging of collections for the museums, Computing and the Arts, and IPCH.
Persistent Linking

In the web environment, change happens rapidly and continuously. Links to resources break frequently as content is moved from server to server. Citations and references that depend upon a link can be compromised if that link is broken. Persistent linking is an important component of any digital content management environment. It allows systems and people to link digital content with the assurance that the link will properly resolve to the content, even if that content has been moved. Many of Yale’s digital services such as the Library’s Orbis electronic catalog, Cross Collection Discovery, and other various digital repositories rely on the Yale Persistent Linking service (YPL) to maintain millions of persistent links pointing to Yale’s valuable digital assets.

YPL is built on CNRI’s Handle System®, a system that has over one thousand installations in more than 44 countries including over 750 installations at universities. Yale’s service is registered with the Handle System Global Handle Registry that, on average, resolves over 68 million requests per month.

Content stewards interested in leveraging this shared infrastructure towards more effective content...
Persistent Links

Broken links often frustrate the exploration of digital content. Persistent links provide a mechanism for ensuring that links to digital content will continue to resolve and provide access to digital content even if the location of the content should change over time. This is done through a service that generates a unique URL and registers a storage location associated to that URL. If the storage location should move, managers register the new location and the service will continue to resolve the initial URL to the new location.

Content Platform
- Digital Preservation
- Persistent Linking

ODAI Statistics
- Yale Persistent Linking Statistics

YDC2 News
- Discover Yale Digital Content
- ODAI announces the Yale Persistent Linking Service

References
- Yale Persistent Linking User Guide
Taxonomy in Drupal

- Vocabulary
  - The big bucket
- Terms
  - Parent/child relationships
- Tags
  - User-defined tagging
Before leaping, consider...

Questions:

- Does your site benefit from taxonomy
- Commitment to ongoing management (new content)
- Technical setup (making taxonomy work the way you want it to work)

- How will you define your taxonomy terms?
- Do you want a single list or do you functionally need parent/child relationships?
- Consider how it will be applied to content types on your website.
CCK vs Taxonomy

- Can't I do this with CCK? Yes, and no...
- How to decide...
- For me, the issue comes down to:
  - Size of the list (management)
  - Application (is the description unique to a single content type, or applicable across multiple content types)
  - How will terms be used on the site?
  - Do I need another dimension of organization?
YDC2 taxonomy, part 2

- We also had a set of references (in the biblio module) that were already tagged with taxonomy terms
- We needed a second facet to arrange these references by a small vocabulary.
- Added a new taxonomy that we could then sort on with a view
- Let's take a look.
How to Add Taxonomy to D7

- Possible Content Type Fields (structure)
  - Artist
  - Album title
  - Songs
  - Artwork
  - Location (built-in field)

- Taxonomy (additional depth and context)
  - Genre
  - Continent

- Let's take a look.
In Summary...

- Make Taxonomy part of the planning process
- Think about implementation
- Think expansively about your site and its mission... how does taxonomy facilitate the user experience.
Beyond Buckets: An Introduction to Drupal Taxonomy

Sarah Prown
prown@reedsparrow.com
Introduction...
I'm a librarian and website consultant based in Guilford. I like to think my background as a professional librarian lets me look at the web development process in an interesting and hopefully useful way. Today, I'm going to share some of the experiences and insights I've had while developing websites (primarily for academic and educational clients), and working with Drupal taxonomy.
What is taxonomy? That seems like a good place to begin.

It is a:
• A focused system of classification
• A tool for organization
• An opportunity for description
• A place of entry (an access point)
• A method for imposing intellectual control
How is taxonomy used?

- Scientific classification
- Libraries (subject headings)
- Your sock drawer?
- Taxonomy is used any place you want to exert intellectual control over a diverse collection of information (or objects, or ideas, or...)

Taxonomy surrounds us.

- Scientific classification
- Libraries (subject headings)
- Your sock drawer, perhaps?

Taxonomy is used any place that you want to exert intellectual control over a diverse collection of information (or objects or ideas or...) You get the picture.
Let's think for a moment how taxonomy is used in a library. Libraries have multiple ways of organizing their collections – author, format, subject... just to name a few. Subject access is typically what we think of when we think taxonomy.

Now imagine a library organized by color! The books about dogs are suddenly much harder to find!
Who remembers this?

The subject catalog. A brilliantly browsable tool for finding all books on a particular topic in a single, convenient drawer. Of course, sometimes you had to visit the miniscule 5-volume list of Library of Congress Subject Headings and wade through a morass of See Also's and Used For's to get to the right term, but a well-referenced subject catalog was just the thing for a last minute research paper...
But then we got computers... and internets. And companies (and libraries) spent zillions of billions of dollars to create systems of intellectual control and access over an exponentially expanding universe of websites. This is Yahoo circa 2002, I believe. It was handy, but it couldn't begin to scale with the internet.
And so we end up with this. No intellectual control, but here's where we often start our interaction with the web. Why?
Thanks for the library lesson, but...

- What does this have to do with my Drupal site?
- Let's take it back a step, and talk about...

**RELEVANCE**

Why does Google work – it's a brilliant tool and the algorithms and paid content generally deliver some relevant content.

And that's where taxonomy comes in – how can we build RELEVANCE into our Drupal websites?
What is a website?

- It's a collection of information/data/etc.
- It's a tool.
- It's an experience.
- Its success or failure is determined by one user's ability to find what they need at the time they need it...

I think it helps to think about what a website is. It's easy to lose sight of this, as the web and websites are so intertwined with our daily lives. But why... what works? What doesn't? What are you trying to achieve with this site, with this page, with this link?
Beyond Buckets

• Think about your favorite websites...
• What works?
  ● Transparency
  ● Discovery
  ● Serendipity
➲ Users invited in to explore, make connections.
Finally, buckets... we're getting somewhere. Content Management Systems encourage bucket thinking – content types, architecture, menus, hierarchies.

These are all essential elements, by the way. Developing a strong internal organization and structure is always the first step in a web development project.
Another dimension...

- Taxonomy allows you to rethink/reorganize content in new ways.
- New connections
- New pathways to content (no orphans!)
- Serendipity and Discovery
- Flattening of hierarchy
- Tag clouds
- User tagging (free-tagging, community)

Taxonomy lets us add another dimension...
You can organize or present your content in new ways. It builds new connections, pathways to content that might otherwise be orphaned on a deep page. It invites serendipity and discovery (remember that browsable card catalog!) and flattens the imperial hierarchy of menu-based browsing (hello google!) Tag clouds and communities offer other dimensions to a website.
Drupal.org uses a combination of taxonomy (categories) and exposed filters to allow users a rich and fast tool for browsing their 20,000-plus modules.
When no category is specified (ANY), we get this broad overview of taxonomy and the most frequently-used modules for each category.
The YDC2 Website

- Challenges:
  - A relatively new department at Yale
  - Marketing (who we are/what we do)
  - Education
  - Communications
- Taxonomy allowed us to address these challenges

In this next example, the website I developed for the Yale Digital Collections Center had several challenges from the outset.

- It was a relatively new department, so how to develop an identity through the website?
- How to communicate the breath and depth of the department and its activities.
- Content-rich site – projects, news, documentation, presentations, and references to professional literature.

Taxonomy allowed us to build a system for addressing these challenges.
On every subpage, there is a sidebar list of taxonomy terms. This carefully collected alphabetical list allows users to see the wide range of elements that come together within the department.
Content Creators can apply one or more of these terms to new or existing content in a number of content types. Not all content types have taxonomy attached on this site.

In this case, taxonomy terms appear unobtrusively under the title, inviting further search and discovery.
When a user clicks on a taxonomy term from the sidebar or within a node, a taxonomy term display appears. This display is built into the drupal core. Custom breadcrumbs display the parent/child relationship in the taxonomy (more on that later...)

Here you see several elements:
• The term and its definition
• All the content on the website that's been tagged with that term, sorted by content type
Drupal taxonomy consists of three parts.

- **Vocabulary** – the big bucket to collect your terms
- **Terms** – the list of words with or without parent/child hierarchy
- **Tags** – the capacity for user-defined tagging of content. Free-tagging. A folks-onomy.
Before leaping, consider...

- Questions:
  - Does your site benefit from taxonomy
  - Commitment to ongoing management (new content)
  - Technical setup (making taxonomy work the way you want it to work)
  - How will you define your taxonomy terms?
  - Do you want a single list or do you functionally need parent/child relationships?
  - Consider how it will be applied to content types on your website.

As you consider whether or not taxonomy is appropriate for your drupal site, here are some things to consider.
CCK vs Taxonomy

- Can't I do this with CCK? Yes, and no...
- How to decide...
- For me, the issue comes down to:
  - Size of the list (management)
  - Application (is the description unique to a single content type, or applicable across multiple content types)
  - How will terms be used on the site?
  - Do I need another dimension of organization?

It's also important to consider whether this job is better served by taxonomy or a CCK field? You can replicate taxonomy on a content type with CCK.

This website has a nice chart about when to use taxonomy and when to use CCK.

For me it comes down to:
- Size of the list (taxonomy management tools are nice)
- Application
- How the terms will be used on the site (navigation? Organization?)
- And most importantly – does this site need another dimension of organization.
YDC2 taxonomy, part 2

- We also had a set of references (in the biblio module) that were already tagged with taxonomy terms
- We needed a second facet to arrange these references by a small vocabulary.
- Added a new taxonomy that we could then sort on with a view
- Let's take a look.
How to Add Taxonomy to D7

- Possible Content Type Fields (structure)
  - Artist
  - Album title
  - Songs
  - Artwork
  - Location (built-in field)

- Taxonomy (additional depth and context)
  - Genre
  - Continent

- Let's take a look.
In Summary...

- Make Taxonomy part of the planning process
- Think about implementation
- Think expansively about your site and its mission... how does taxonomy facilitate the user experience.
Beyond Buckets: An Introduction to Drupal Taxonomy

Sarah Prown
prown@reedsparrow.com