Testing Taxonomies

Heather Hedden
Hedden Information Management

Taxonomy Boot Camp, Washington, DC

November 5, 2013

About Heather Hedden

- Independent taxonomy consultant, Hedden Information Management

- Continuing education online workshop instructor, Simmons College Graduate School of Library and Information Science

- Author of *The Accidental Taxonomist* (Information Today, Inc., 2010)

- Previously
  - taxonomy consultant employed by a consulting firm
  - taxonomy manager
  - taxonomist for enterprise search tool vendor
  - controlled vocabulary editor at a library database vendor (in Vocabulary and Quality Management department)

Outline

- Taxonomy testing overview
- Types of tests for taxonomies
  - Card sorting
  - A/B testing
  - User/use case testing
  - Indexing testing
  - QA testing
- Conclusions

Taxonomy Testing Overview

- Taxonomies serve a purpose, and that purpose should be tested.

- All taxonomies, whether created by subject matter experts or taxonomists, should be tested.

- Testing involves participants, as sample or representative users.

- Testing can be simple and basic, or elaborate and thorough, depending on budget.

- Different types of tests are appropriate for different stages of taxonomy development.

- An inappropriate test or inappropriately timed text can be a waste of time and money.

Taxonomy Testing Overview

Different tests for different stages of taxonomy development

- Design and development phase: *to test ideas*

  - Card sorting
  - A/B testing

- Draft completion phase: *to test usability/functionality*
  (may also be considered "validation")

  - User/use case content retrieval testing
  - Content indexing testing

- Implemented taxonomy: *to periodically test quality*

  - QA testing

Card Sorting

- Method common in information architecture for website menu label organization

- Term names/label/topics are written down each one to a card, and the cards can be sorted into groups.

- Traditionally done with actual index cards. Now usually done through software, usually drag-and-drop and online to allow remote access.

- Involves participation of multiple stakeholders or test-user subjects

- Two types:

  1. **Open sort**: participants group terms and assign the groups category names of their own choosing

  2. **Closed sort**: Categories are pre-defined, and participants place terms in the appropriate categories

Subscription-based web services:

www.optimalsort.com

http://uxpunk.com/websort/ (formerly known as Websort)

OptimalSort

Purchase a cell phone online

A tool to calculate the best cell phone plan for me

Internet connection speed test

International calling rates for my home phone

Add-on services for my home phone

Pay my bill online

Accessories for my cell phone

BananaCom's contact phone number

Downloadable games for my cell phone

The date that my next home phone and internet bill is due

What to do when my cell phone has been broken or stolen

Career opportunities at BananaCom

Deals for home internet and phone bundles

Change my home internet plan

Step 1

Take a quick look at the list of items to the left.

We'd like you to sort them into groups that make sense to you.

There is no right or wrong answer, just do what comes naturally.

Step 2

When you're ready, drag an item from the left to create your first group.

▼ Click to rename ✕

Purchase a cell phone online

Step 3

If you already know what you'd like to call this group, click the title to rename it now. If not, you can do this later.

Step 4

Now you can either put more items into this group, or make more groups. It's up to you!

When you're done, click "Finished" (top right).

▼ **Products and Services** ✕

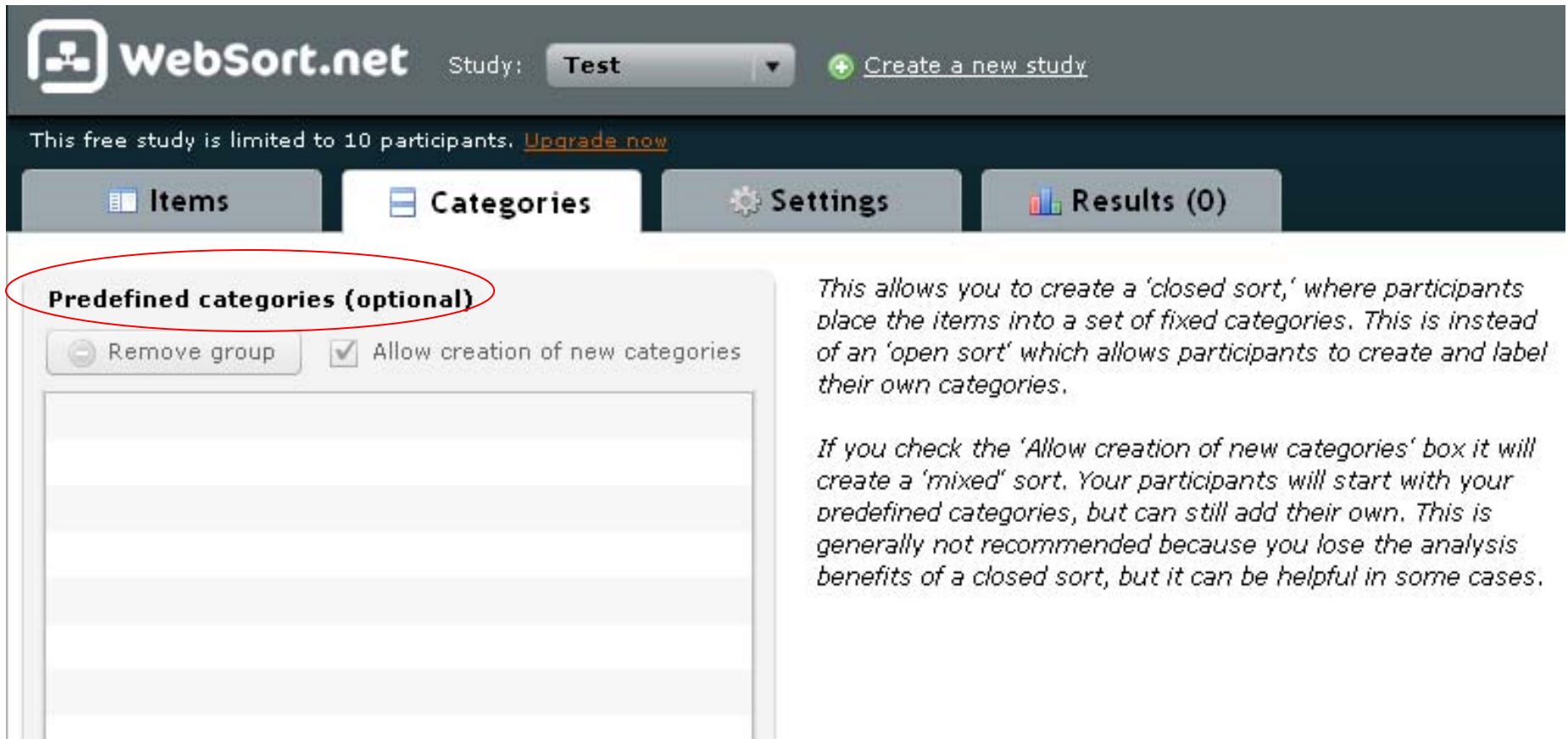Add-on services for my home phone

Discounted cell phones

Purchase a cell phone online

Drop to create group

International calling rates for my home phone

Open Sort

formation Management

Card Sorting

Closed Sort

Card Sorting

Although sometimes called "card sort testing" this is for testing *ideas* for taxonomy development, not for testing taxonomy functionality.

- Open sort - early in the taxonomy development process
- Closed sort - later in the process

Additional, functionality testing is still needed.

Subtle differences in wording, such as part of speech, can effect sorting choices.

Example: Same or different categories?

Scuba diving

Racquet ball courts

Aerobics classes

Consider changing term names to make them similar in a group.

Card Sorting – Issues

More suited for website navigation design than for taxonomies:

- **Open card sort** – rarely needed for taxonomies

  - Unlike website navigation menus, taxonomies are rarely created from scratch. Usually there are top categories to start from.

- **Closed card sort** – purpose of finding the desirable broader category is often not needed

  - Unlike website navigation menus, a topic may have more than one broader category (polyhierarchy) in a taxonomy.

- Card sorting is useful for a hierarchy of only one narrower level. Impractical to test multi-level hierarchies, which require multiple tests

- Card sorting is no useful for faceted taxonomies.

A/B Testing

- Test subjects are people who will use the taxonomy to find content.

- Test-users are presented with two different possible scenarios ("A" and "B") and asked which they prefer

- Like closed card sort, done at a similar stage to test taxonomy ideas

For taxonomy A/B testing, alternative could be:
➢ Different wordings of category labels
➢ Different ordering/arrangement of categories
➢ Different location of subcategories

- Often done in a user interface mock-up / web page wireframe.
- Can be performed any time in the taxonomy design and build process.
- Useful when undertaking a taxonomy redesign.

A/B Testing – Example

Do you like A?

(logically grouped)   OR

Do you like B?

(more revealed at top level)

Shop Cameras & Camcorders

Digital Cameras

Film Cameras & Accessories

Camcorders

Camcorder Accessories

Memory Cards & Readers

Photo Printers

Digital Photo Frames

Binoculars, Telescopes & Optics

Digital Camera Accessories

Online Photo & Video Services

Camera & Camcorder Services

Shop Cameras & Camcorders

Digital Cameras

Digital SLR Cameras

Compact System Cameras

Point & Shoot Cameras

Film Cameras & Accessories

Camcorders

Camcorder Accessories

Memory Cards & Readers

Photo Printers

Digital Photo Frames

Binoculars, Telescopes & Optics

Digital Camera Accessories

Online Photo & Video Services

Camera & Camcorder Services

A/B Testing – Issues

- Most suitable to compare proposed top-level categories

- Not practical to conduct a detailed term-by-term comparison

- Most effective if making use of graphical user interface design

  - Existing or proposed design can be altered in a drawing program

- Test users may not have the time or patience for numerous A/B tests.

- Need to decide what to compare and how many comparison tests to make, to conserve time and resources.

User/Use Case Testing

- A form of taxonomy validation

- Tests to see if taxonomy will perform as hoped in search/retrieval

- Test-users are people who will use the taxonomy to find content.

  - Particularly applicable for users of internal/enterprise taxonomies

  - For public/subscriber access to taxonomies, instead of actual users, subject matter experts or customer support may suggest "typical" use cases.

- Can test deep hierarchical or faceted taxonomies

User/Use Case Testing – Procedure

1. Test-users are asked to prepare several use cases (information seeking scenarios), through interviews or written descriptions

   ➢ Most are "typical" scenarios
   ➢ One or two may be recent challenging scenarios

2. Test-users are asked to browse the draft offline taxonomy to look for terms under which the content for each scenario might be found

   Test users perform the test, either:

   a) in the test administrator's (taxonomist's) physical presence
   b) via screen-sharing with verbal narration explaining choices made
   c) independently in an offline worksheet, and then reviewed in a verbal meeting

3. Test administrator takes notes regarding problems in finding taxonomy terms for the use cases.

User/Use Case Testing – Examples

Use case scenarios are initially narrative

- For licensed content subscribers:

    Jeff has a question about which employees, within the bank, meet the SAFE Act definition of a "loan originator" and therefore would be required to register.

- For internal web content management

    A staff member of the Content Group would like to find a seasonal banner ad for a specific brand to upload it.

    Minimally must answer: who, what (which can be complex), and for what purpose

User/Use Case Testing – Outcomes

Noted findability problems should be considered as indications for:

- Additional taxonomy terms

Or if the terms exist:

- Additional nonpreferred terms (synonyms) to point to existing terms
- More polyhierarchy (multiple drill-down paths to the same specific term)
- More associative (*See also*) relationships (if supported)

to help guide the user to find the desired concepts

User/Use Case Testing – Issues

- Too complex for non-employee test subjects recruited from the general public (unlike card-sorting and A/B testing).

- Requires advance planning and preparation

  - Test users may have difficulty formulating use cases

  - Taxonomist should develop use cases out of stakeholder interviews

- Important use cases could still be overlooked.

Indexing Testing

- A form of taxonomy validation

- Tests to see if taxonomy is suitable to index/tag/classify intended content

- For manual indexing, test-users are the indexers or content creators who assign the tags/categories to content.

- For auto-classification, taxonomist tests indexing.

  - If implemented in the system, conduct testing in the system

  - If not yet implemented in the system, conduct testing in a similar manner as for manual indexing.
    (Additional testing, once implemented, will still be needed.)

- Can test deep hierarchical or faceted taxonomies

Indexing Testing – Procedure

1. Test-user indexers, content creators, or stakeholders identify a set of varied sample documents/content assets that need indexing.

2. Test-indexers are asked to browse the draft offline taxonomy to look for terms to classify/tag each document or content asset.
    (Usually done independently)

3. Taxonomist does the same.

4. A meeting between the taxonomist and the test-users discusses the choices of indexing terms made

    ➢ separate meetings, if each test their own individual set of documents

    ➢ joint meeting if the same content was test-indexed by all participants

Indexing Testing – Example worksheets (Faceted taxonomy)

	A	B	C	D	E	F	G	
1	Document	Category > subcategory	(optional) Additional Category >	Source/Author	Language	Region	Brand	Content 1
2	2010_BROCHURE_PRICING							
3	Accounting.htm							
4	Airline_Market_Information							
5	alliance_introduction_letter							
6	asia_pacific_sales_overview.pdf							
7	Copy of QA_Self_Inspection_Form							
8	core_brands_site_inquiry_reports							
9	emea_accounts_by_account							
10	europe_middle_east_africa_training_curriculum_and_courses							
11	Getting_Tour_Market_Business							
12	GO_GLOBAL_PROFILE							
13	gso_government_Terminology_Websites							
14	gso_transient_Commercial_Corporate_Traveler							
15	gso_US_military_base_links							
16	Guides and FAQ on the Opaque program							
17	How to get more AAA business							

A	B	C	D	E	F	G
Content Item/Asset	Owner of Content	Language of Content	Image Type	Promotion Type	Location	Brand
ECPAT.png	Legal	English	Icon		N/A	All Brands
explore_turkey_header	Online Retailing	N/A	Landing Page Header Image, G	Destinational Promot	Turkey	▨▨▨ Hotel Group
explore_turkey_metapath.xhtml	Online Retailing	English	[none]	Destinational Promot	Turkey	▨▨▨ Hotel Group
MT_2013_Summer_header_Stay3	Online Retailing	English	Landing Page Header Image, S	Multiple Night Discou	N/A	▨▨▨
MT_2013_Summer_slider_Stay3	Online Retailing	English	Slider Image, Stock or Lifestyle	Multiple Night Discou	N/A	▨▨▨
MT_2013_Summer_Stay3_140x102	Online Retailing	N/A	Thumbnail, Stock or Lifestyle		N/A	▨▨▨
print-logo.png	Brand Marketing	English	Brand Logo			▨▨▨
QR LP hotel listing _ME and Turkey.xls	Online Retailing	English	[none]			

21

Indexing Testing – Outcomes

- Inability to find a taxonomy term may indicate the need for...

  - Additional taxonomy terms
  - Additional nonpreferred terms (synonyms) to point to existing terms
  - More polyhierarchy (multiple drill-down paths to the same specific term)
  - More associative (*See also*) relationships (if supported)

- Uncertainties or inconsistencies may indicate that...

  - a taxonomy term should be reworded for clarity or nonpreferred terms need to be added (usually the case for dedicated indexers)

  - The vocabulary size is larger than it needs to be and could be shortened and simplified (more often the case for content authors for whom indexing is an undesired task)

- Use of N/A may indicate that an indexing policy should not require the use of a certain facet.

Indexing Testing – Issues

- Task is time-consuming: sufficient participation of enough participants with a significant number of documents can be difficult

  - ➢ Aim for 4-5 documents/assets per participant

- Not simulating the indexing/tagging user interface

- Inconclusive results, that may just require policies and training

QA Testing

- Test searches (or browsing) to test taxonomy in retrieval results.

- Performed after the taxonomy and content is implemented in the system and periodically thereafter.

- Taxonomist performs random searches

  ➢ Selects random terms from the taxonomy to see if appropriate content is retrieved (to test precision)

  ➢ Identifies content items, and checks to see if appropriate taxonomy terms retrieve it (to test recall)

Issues:

- Usually cannot discern if errors are due to taxonomy problems or indexing problems. Need to investigate both.

- Not always easy to test recall.

© 2013 Hedden Information Management

Conclusions

Test	Testers	Purpose
Card sorting	Stakeholders or sample users/searchers	To test taxonomy structure ideas - but often not practical for taxonomies (in contrast to testing website menu navigation)
A/B Testing	Sample users/searchers	
Use case Testing	Sample users/searchers	To test suitability to retrieve desired content
Indexing testing	Sample indexers/taggers or taxonomist	To test suitability to tag content and appropriateness for content scope
QA testing	Taxonomist	To test continued precision and recall in retrieval results

Conclusions

Evaluating vs. Testing Taxonomies

- Evaluating a taxonomy
    - to determine if it's well designed and constructed
    - does not always require having sample content or sample users
    - may be done on an existing or implemented taxonomy
    - done by an expert taxonomist

- Testing a taxonomy
    - focuses on the specific application and use of the taxonomy
    - involves using sample content and sample users
    - part of the taxonomy development process
    - done by sample or representative users, facilitated by a taxonomist

Testing (aside from idea-testing in card-sorting and A/B testing) should come *after* initial taxonomy evaluation and revisions.

Questions/Contact

Heather Hedden

Hedden Information Management

Carlisle, MA

heather@hedden.net

978-467-5195

www.hedden-information.com

www.linkedin.com/in/hedden

twitter.com/hhedden

accidental-taxonomist.blogspot.com

© 2013 Hedden Information Management