Managing taxonomies strategically

What are taxonomies?

Taxonomies are structures that provide a way of classifying things -- living organisms, products, books -- into a series of hierarchical groups to make them easier to identify, study, or locate. Taxonomies consist of two parts -- structures and applications. Structures consist of the categories (or terms) themselves and the relationships that link them together. Applications are the navigation tools available to help users find information.

Taxonomy structures

A taxonomy structure might look something like this:

```
Industries
    Financial services
    Manufacturing
    Retail
    Transportation see also "Employment [transportation]"
        Air
        Ground transport
        Maritime
        Space
Ocean transportation see "Maritime transportation"
```

Taxonomy structures typically have the following elements:

• List of standard terms (in the example above, we've used "maritime" instead of "ocean" as the standard term)

• Hierarchical relationships ("Transportation" is subordinate to "Industries")

• Cross references (If you're looking for recruiters specializing in the transportation industry, look under the category "Employment [transportation]." You may call shipping via boat "sea transportation," but our standard term is "Maritime transportation")

Taxonomy applications

Taxonomy structures can be used in a variety of applications, such as helping:

• researchers find source materials
• readers locate information in a book
• Web visitors locate information in an electronic journal
• buyers locate products and services
• decision makers locate sources of expertise

Taxonomy structures can also make automated processes more efficient. For example, taxonomy terms can be used...

...in a search engine query to help users find information more easily;
...in a filtering program to personalized e-mail alerts or Web sites;

Taxonomy applications have three key elements -- people, tasks, and sources (or content). Note that the same taxonomy structure can serve multiple applications and can be applied to different kinds of content (i.e. articles, books, videos, speeches).

What do taxonomies do?

Taxonomy structures and applications working together can help people perform three kinds of tasks:

1. *Identification* -- Is this a service we can provide? Is there a product that will do this? Have we solved this problem before? Does this type of information exist?

2. *Discovery* -- What's related to the issue or problem at hand? Who's working on it?

3. *Delivery* -- If something is available, where or how do I get it?

A good back-of-the-book index does all three tasks through an alphabetical list of terms, cross references to related terms, and page references. Roget's Thesaurus, on the other hand, does only the first two tasks. It tells you that a word exists and directs you to related terms, but it doesn't link the terms to a source.

Taxonomies and trust

In addition to performing the basic three functions, a taxonomy should also inspire trust. The user should feel confident that the taxonomy will help him find the information he seeks -- if it exists. Unless a local business is very new, very small, or very specialized, you can be pretty sure you'll find it in the local yellow pages. But you can't expect to find all the businesses in the world on Yahoo's Web site -- even though you can use Yahoo's search feature to find very small and very specialized firms in some remote locations. And there's the rub. As more information gets into electronic format and becomes available over global networks, it gets harder to ensure that any one taxonomy is both sufficiently specific and comprehensive.

Eccentric taxonomies

Although it is possible for one taxonomy structure to serve multiple applications, in practice most taxonomies are eccentric -- suitable for only one environment and application. Most business taxonomies are highly customized, and it's not unusual for a single company to use multiple taxonomies for different functions or applications (e.g. one for marketing, another for product development). In fact, one could argue that your taxonomy -- how you look at the world -- is an important source of strategic advantage.

Managing business taxonomies

The popularity of Web publishing has generated a demand for classification tools and techniques traditionally available only to libraries and large commercial publishers. At the same time, global companies are struggling to integrate hundreds, sometimes thousands, of "eccentric" taxonomies

used by different geographic regions, corporate departments, and business units. The Web is empowering small publishers to create sophisticated electronic publications in the same way that desktop publishing empowered them to produce print publications twenty years ago. It is creating new demands on corporate information managers, who suddenly need to acquire skills in cataloging, acquiring, aggregating, and syndicating content.

How to create a business taxonomy

A taxonomy usually starts with a list of standard terms -- a "vocabulary" -- that describes the content to be organized. The Montague Institute vocabulary grew out of the index terms used in the first of our briefings (for a list of our vocabulary terms, see the A - Z Index on the Indexes and thesaurus page).

The second step is to add relationships among the terms. Relationships include cross references from nonstandard terms (e.g. FASB) to standard terms (e.g. "Financial Accounting Standards Board"), from narrower terms to broader terms (e.g. "transportation" see

also "industries"), and from one term to a related term (e.g. "indexing" see also "taxonomy"). Many people also incorporate definitions and notes (e.g. "mining" refers to a computer data extraction technique, not a method of extracting ore from under the ground). This part of the taxonomy is often called the "thesaurus." Like the familiar Roget's Thesaurus, it contains synonyms, but is also does a lot more.

The vocabulary and thesaurus constitute the taxonomy structure. The next step is to connect the terms with sources -- Web sites, documents, people, or pages in a book. Typically this happens in the taxonomy application, along with sorting and formatting the terms.

Human vs. computer taxonomies

Although most taxonomies are created by humans ("indexers" in the publishing world, "catalogers" in the library world), it is possible to create taxonomies automatically using artificial intelligence programs. These programs are most cost effective for very large collections, such as millions of news

Montague Institute teaching lab showing the TERMS (vocabulary) section

stories, thousands of crime reports, or large corporate repositories. However, even when computer programs are used, they almost always need editing to accurately reflect your specific content. For example, the programs have to be "taught" that when people use "printer" as a search term, they mean a device that produces hard copy of a computer file, not a person that operates a printing press or a commercial printing firm. "Training" a taxonomy program may involve changing parameters or editing a "draft" taxonomy produced by the program.

Managing taxonomies strategically

Managing taxonomy as a service requires three adjustments:

1. Remember that it's more important to serve the real needs of users than it is to produce an "ideal" textbook taxonomy. The taxonomy structure itself is necessary, but don't neglect assistance in application development, training for editors, authors, and users, and tools for authoring and maintenance.

2. Educate content owners and editors about the real costs and benefits of taxonomy products and services.

3. Focus more on the categories that customers use and the places they look for information. This means using layman's terms instead of technical terms, using a thesaurus for cross references, and including people and informal communications (i.e. e-mail, discussions) as sources.

Conclusion

A growing number of content owners and business unit managers are beginning to realize the value of taxonomies to improve access to information, leverage existing content to create new products and services, and facilitate electronic commerce. Although taxonomies are not new, they need to be integrated and customized for a variety of business applications. As information professionals, it's important that we offer content owners taxonomy services that are both convenient and cost effec-tive. At the same time, we need to link specialized taxonomies together and leverage investments in taxonomy infrastructure across the entire enterprise. If we're successful, taxonomies will realize their potential as strategic assets.