Exploring the Taxonomy Underlying www.plainlanguage.gov
by John Heffernan

As part of the plain language project we paid special attention to the taxonomy underlying the site structure. In general, site taxonomy begins with discovering the most important, broad categories of information discussed on the site. As we began to identify topics of discourse, we began to see a preferred vocabulary emerge to form conceptual chunks, or facets, of site knowledge. For the plain language site, we identified four facets: resources, law implementation, training and writing guidelines/examples.

Within each facet, we envisioned how the vocabulary could be organized into logical hierarchies, reflecting the organic patterns of thinking of the community of site users. Strictly speaking, each hierarchical facet is itself a taxonomy. (By definition, a taxonomy is a hierarchical organization of categories.)

We also applied principles of vocabulary control so each term expresses a clear, ambiguous meaning, while reflecting the natural intermixing of ideas. Some of the most important vocabulary principles include the following:

a) Uniqueness (terms should appear once)
b) Terseness (three words is getting verbose)
c) Polyhierarchy (ideas can belong to different ways of thinking or be expressed in different context. For example, tuberculosis can be a disease of both the lung and of the bone, so lung diseases and bone diseases can properly reference tuberculosis).

Establishing a taxonomy is one thing; embedding it into site architecture is another. We discovered that the plain language site has a two-level taxonomy structure, each relating differently to site content and use. One taxonomy applies to navigational labels and to page content organization and language. The other taxonomy provides terse metadata for leveraging the structure of the clear writing guidelines in order to access the example content. A multi-layered taxonomy, or even use of multiple taxonomies in a single site, is not uncommon, given the need, as in this case, to provide a taxonomy for things and another taxonomy for properties and affects.

The navigational labels and relationships enable users to access the requirements of the federal guidelines document and help them know legal standards for complying with these requirements. They support people in learning more about plain language and finding training and resources related to clear writing. These labels also can support advocates of plain language and can help people to learn more about benefits and uses.

On the other hand, the writing examples metadata proved to be more challenging. The current site has a unique taxonomy where people can find examples via groupings such as by agency. But the users strongly voiced their need to find examples based on new federal guidelines for clear writing. We realized this difference was one we needed to address – but also saw that it could serve as a stand-alone project. We intend to explore this challenge in a later class.

For the current site, we focused on uniqueness of terms and on the framework language for page content. We realized people were having challenges with similar words such as guidance and guidelines so we developed labels that supported these nouns (such as writing guidelines). We approached the terseness of labels as part of our optimizing process – we asked how we could focus the label on what people want to do and how they might search for a term. We incorporated polyhierarchical thinking by using checklists and other routes back to content-rich pages. ■

John Heffernan has more than 30 years experience in publishing, leading teams in the creation and production of scholarly and legal research collections, metadata-based finding-aids and indexes. He has worked with development teams, customers and end-users to test, deploy, implement and technically support web and enterprise workflow solutions for scientific, technical and medical journal peer review and for large-scale print production processes. He can be reached at john.heffernan@restonindexing.com

Resources Mentioned in the Article