Keywording Use Cases

**Goal: to increase the use of terms in a controlled vocabulary by providing tools that aid metadata creators in identifying and incorporating suitable terms**

The terms to be incorporated may come from an LTER-wide vocabulary or from a site-specific controlled vocabulary.

**Use case 1: Browse or search for terms**

Steps: In a web form or stand-alone application

1. Browse or search (with autocomplete) for preferred terms
2. Check off desired terms for addition to metadata
3. As things are checked off, an internal list is generated
4. When selections are complete, “Finish” button is hit
5. The application returns either:
   1. An XML snippet suitable for input into an Ecological Metadata Language (EML) document or a web service client
   2. A plain-text list

**Use case 2: Case 2 is a Drupal-specific version of use case 1**

Steps:

1. Populate Drupal web site with polytaxonomy
2. Within the Drupal Metadata Editor
   1. Browse a drop down list of levels, or
   2. search to find terms
3. Select term you want and it is automatically added to backend database that is used by the module that creates EML within Drupal

**Use case 3: Semi-automated keywording**

Steps:

1. Identify (URL) or upload draft EML document
2. Use the HIVE web service developed by Duane Costa to suggest probable preferred terms
3. Check off suitable terms (including “select all”)
4. Results can be returned as:
   1. A plain-text list with terms listed one per line for cut-and-paste,
   2. An EML snippet to screen, for cut and paste into a document,
   3. A revised EML document with selected terms added, or
   4. An XML document with keywords (in a keywordset node, including thesaurus) to be used with a web service client (thus allowing additions to relation databases etc.)

**Additional Considerations**

* Although principally designed for keywording of dataset metadata, use cases 1 & 3 could also be applied to other applications such as keywording of documents or projects.
* Use of web services or standard exchange formats (SKOS) will make it relatively easy to use a variety of lexical resources (e.g., site-specific vocabularies)

**Use case 4: Best Practices**

Effective use of controlled vocabularies depends on more than suitable tools. Individuals adding terms to metadata documents would also benefit from a “best practices” guide that addresses use of controlled vocabularies, with the goal of assuring that LTER data is discoverable.

Some examples of what might be included are:

* Use the most specific terms you can
* Specify how many or what categories of terms should be included where applicable - examples
* Specifing a desirable number of terms
  + E.g., At least one term from at least X of the LTER taxonomys
* Should have at least one core area

**Use case 5: Rating Documents**

Tools that identify metadata documents which would benefit from additional keywords from the controlled vocabulary can be used to check for conformance with best practices.

Steps:

1. Run document through congruency checker
2. Receive a count of how many keywords and taxonomys are represented in an EML document