

Best Practice Poster: **Utilizing Drupal for the Implementation of a Dublin Core-Based Data Catalog**

Lisa Federer
National Institutes of Health Library,
United States
lisa.federer@nih.gov

Keywords: NIH; National Institutes of Health; data catalog; Drupal; content management system.

1. Objective

To create a data catalog suitable for use within the context of biomedical and health sciences research. The ideal catalog would allow researchers to easily describe their data using Dublin Core Metadata Terms and subject-appropriate controlled vocabularies, as well as provide search and browse capabilities for end users to enable data discovery and facilitate re-use.

2. Setting

The National Institutes of Health (NIH) Library serves the community of NIH intramural researchers, which includes over 1,200 principal investigators and 4,000 postdoctoral fellows conducting basic, translational, and clinical research on its primary campus in Bethesda, MD, and several satellite campuses.

3. Methods

Drupal, a free and open-source content management system, was utilized as a framework for a data catalog using the Dublin Core Metadata Terms. Using the Structure function within Drupal, the research data informationist at the NIH Library constructed a pilot system that utilized Dublin Core Metadata schema and relevant biomedical taxonomies. This pilot system can be adapted to the needs of a variety of basic, translational, and clinical research applications.

4. Results

The pilot system is currently undergoing testing with researchers within the NIH intramural community. Results will be available by the time of the DCMI 2014 conference.

5. Conclusions

A data catalog that utilizes an extensible metadata schema like Dublin Core and an open-source framework like Drupal provides users a powerful yet uncomplicated method for describing their data.

5. Implications

As funders and publishers increasingly require data sharing, researchers will need simple, intuitive methods for describing their data. Open-source systems like Drupal and extensible metadata schema like Dublin Core will likely play a large role in data description, thus making data more discoverable and facilitating data re-use.