

Evaluating Vocabulary as a Component of Faceted Search

Ellen Roscoe Iverson, Cathryn A. Manduca, Sean FOX Science Education Resource Center, Carleton College



The Science Education Resource Center (serc.carleton.edu) site provides a portal for cross-cutting geoscience topics that unites several NSDL digital libraries. To support exploration of these collections we are developing several controlled vocabularies. Web resources are tagged with terms from these vocabularies allowing us to provide a faceted search through the topic with easy exploration along axes of interest. This poster highlights the methodologies used to evaluate the vocabularies across these collections. The collections of digital libraries include NSDL projects such as Starting Point: Teaching Entry Level Geosciences, Earth Exploration Toolbook, and Microbial Life Educational Resources, as well as DLESE Community Services: Using Data in the Classroom, Teaching Quantitative Skills, Integrating Research and Education and Preparing Teachers to Teach Earth Science.

Controlled Vocabularies for Specialized Collections

Terms in these vocabularies need to:

- ❑ have a meaning that educators will immediately grasp: be evocative and resonant
- ❑ highlight ideas of greatest interest to educators.

Evaluation methodologies

Closed card sort

- Earth System Science vocabulary: map teaching activities to existing vocabulary.
- Using Data vocabulary: rate correspondence of teaching activity to vocabulary across multiple vocabulary elements (topic, audience, inquiry level, special interests).

Freelisting

Quantitative Skills vocabulary: identify what quantitative skills in the geoscience domain are most critical.

Open card sort

Quantitative Skills vocabulary: construct vocabulary for existing teaching activities.

Initial Evaluation results

- ❑ **Earth System Science vocabulary closely matches the language used by geoscientists.** 27 of 33 (81%) participants mapped the teaching activity to the assigned term. When asked to rate (on a 1 to 10 scale) how closely a set of vocabulary matched a teaching activity, participants again matched the assigned vocabulary very closely. The assigned vocabulary terms average rating was 8.3 on a 10 point scale.
- ❑ **Geoscience faculty think of quantitative skills differently from faculty in mathematics.** However, the vocabulary of geoscientist appears consistent within their domain. We are using these results based on salience and frequency to shape the existing vocabulary.
- ❑ **Card-sorting and freelisting methodologies fail to identify the use patterns of multiple vertices of vocabulary.** For the Using Data in the Classroom site, no single type of vocabulary or pattern of vocabulary usage emerged. We plan to use focus group and think aloud walkthrough protocol to better understand how the different vocabularies are used.

Quantitative Skills

Basic Skills
Geometry and Trigonometry
Graphs
Functions
Calculus and Analysis
Linear Mathematics
Probability and Statistics
Interdisciplinary Concepts
Technology
Higher Order Skills

Earth System Science

Atmosphere
Biosphere
Climate
Earth surface
Energy/Material cycles
Human
Dimensions/Resources
Hydrosphere/Cryosphere
Ocean
Solar system
Solid Earth
Time/Earth History

Faceted Search

Resources in SERC specialized collections can be explored through a *faceted search* interface. The collection is categorized into multiple *facets*--each controlled vocabulary providing refinement along a different axis. Browsing through these *facets* (vocabularies) can be freely combined with text-based search, allowing users to control their view of the collection.

A screenshot of the SERC faceted search interface. At the top, there is a search bar with the word "pollution" entered and a "Search" button. Below the search bar, the results are displayed in a faceted manner. On the left, there is a "Text Search" label with a red arrow pointing to the search bar. On the right, there is a "Vocabulary-based Hierarchical Browse" label with a red arrow pointing to the faceted results. The results are organized into several sections: "Tools and Data Sources", "Topics", "Ease of Use", "Resource Type: Datasets and Tools", and "Data Source". Each section has a "Show all Examples" link and a "Show All Ease of Use" link. The "Topics" section includes "Atmosphere 1 match", "Human Dimensions/Resources 2 matches", and "Hydrosphere/Cryosphere 1 match". The "Resource Type: Datasets and Tools" section includes "Images, Maps, Animations 2 matches", "Datasets 2 matches", and "Datasets with Tools 1 match". The "Data Source" section includes "Observational Data 2 matches" and "Real-Time Data 1 match". A red arrow labeled "Feedback on current search provided within the vocabulary lists" points to the "Data Source" section.

The user can refine the search results easily along multiple axes--through the provided vocabularies as well as text search. The controlled vocabularies serve double-duty--a browse interface and a feedback tool for search scope.

References

- Fox, S., Manduca, C.M., Iverson, E. Building Education Portals atop Digital Libraries. *D-Lib Magazine*, January, 2005. Available at <http://www.dlib.org/dlib/january05/fox/01fox.html>.
- Sutton, Stuart A. Conceptual design and deployment of a metadata framework for educational resources on the internet *Journal of the American Society for Information Science*, Vol. 50, Issue 13, pp 1182 – 1192.
- Fuccella, J. Using User Centered Design Methods to Create and Design Usable Web Sites, *SIGDOC 1997 Proceedings*.
- Quinlan, M. Considerations for Collecting Freelists in the Field: Examples from Ethobotany, *Field Methods*, Vol. 17, No. 3, pp 219-234.