

## Prospects for integration of Thesauri and Taxonomies across agencies.

Summary of presentation, CENDI staff workshop on Knowledge Organization Systems: Current and Future Uses Among CENDI Agencies, September 16, 2004.

David E. Wojick. Ph.D.  
DOE Office of Scientific and Technical Integration  
WojickD@osti.gov

Most CENDI agencies have thesauri and taxonomies that have been developed for internal purposes, such as indexing and abstracting technical literature. However, broader uses of these tools are now being explored, especially as web-based search aides. Given that science and technology research often cuts across agencies, this search tool development suggests the need for exploring how agency tools might be better integrated. Science.gov in particular might benefit from such an approach.

As examples of candidates for cross agency integration, two experimental approaches ongoing at OSTI were presented, the first by myself and the second by my colleague William Watson. They are part of a project called Word Web.

The first experiment involves using the 200,000+ related term, or RT, relations found in the Energy Subject Thesaurus. These RT relations are mapped using the principle that research communities are language communities. We are exploring the use of such maps to see the underlying concept families of, and between, various scientific and technical communities.

Two such maps were presented. The first is called the Atoms map and entitled "From Neutrons to Nuclear Power: shared knowledge enables progress." It uses RT relations to show potential knowledge flows from neutron research to site selection for new nuclear power plants.

The second map is called the NMR RT-Community map and is entitled "A Related Term (RT) Family for Nuclear Magnetic Resonance." This map shows part of the cluster of concepts used by the NMR research community. It is believed that distinct research communities can be distinguished using this kind of semantic mapping. We are also exploring 3D visualization of RT structures for this kind of flow and community discovery.

William Watson presented the second experiment, a prototype search tool that uses a new Energy Taxonomy that he has developed, in order to find search terms in the Energy Thesaurus. This taxonomy has a powerful 3 dimensional structure, combining disciplines, types, and Thesaurus terms. See our preprint -- "A Taxonomy of Energy Subjects briefly explained" -- by David Wojick, PhD, & William Watson, PhD.

Neither of these approaches -- the RT mapping techniques, nor the Energy Taxonomy - are agency specific. Each presents an approach that could be used to integrate

existing agency Taxonomies and Thesauri, especially as search tools. No doubt there are other approaches that are also worth exploring.

David Wojick