Fall 2005

01001100 01010100 01000101 01010010
LTER DataBits
Information Management Newsletter of
The Long Term Ecological Research Network
01001100 01010100 01000101 01010010

Good Reads

Incorporating Semantics in Scientific Workflow Authoring

- David Ribes (PAL/CCE)

C. Berkley, S. Bowers, M.B. Jones, B. Ludascher, M. Schildhauer, J. Tao. Incorporating Semantics in Scientific Workflow Authoring. Proceedings of Scientific and Statistical Database Management, SSDBM'05, Santa Barbara, CA.

This article presented at the SSDBM (http://2005.ssdbm.org/program.html) and available on the LTER Information Manager Meeting website (http://gce-lter.marsci.uga.edu/lter_im/2005/app/resources.asp?webpage=references) describes an instance of the merging of ontologies with workflow systems in the with a particular workflow system called Kepler being used by the SEEK (Science Environment for Ecological Knowledge; http://seek.ecoinformatics.org/) together with a number of partners. Kepler is a promising research avenue, a tool designed with a generalized support for scientific workflows which are locally tailored for specific tasks through ontology-enabled, domain-specific customization. Workflows are the automation of scientific processes related to data manipulation and representation, while ontologies are system accessible scientific terminologies and computable linkages between them. This article describes how in combining ontologies and workflows a system for data discovery, manipulation and representation was created which is both domain specific and cross-disciplinarily configurable. The authors briefly describe problems in creating workflows and ontologies, and the future possibilities of automated data integration through re-use of semantically annotated workflows.

The Meaning of Everything

- Lynn Yarmey (PAL/CCE), Karen Baker (PAL/CCE)

Winchester, Simon. The Meaning of Everything: The Story of the Oxford English Dictionary. Oxford: Oxford University Press, 2003.

Pearl S. Buck noted that 'One faces the future with one's past.' As the LTER community plans for the future of data accessibility and interoperability, we may draw upon the experiences provided by an example from our recent past. In his book, The Meaning of Everything: The Story of the Oxford English Dictionary, Simon Winchester sends a quiet word of encouragement to those on the designing edge of data organization requirements. Our own metadata processes seem to have so much in common with the story he tells of the creation of the infamous English dictionary.

The OED began as an attempt by the Unregistered Words Committee of the Philological Society (founded in 1842) to mend and improve upon the lesser dictionaries of the age, though not long into the endeavor the need for a complete summation of the language became apparent. With supporting institutions and the editors themselves repeatedly underestimating the enormity and daunting complexity of the project, funding was a constant issue despite the collaborative network of volunteers numbering in the hundreds. The philosophical balances - between short and long-term goals, between meeting a budget and maintaining excellence, between highly specific definitions and the very general, and between creating a fixed standard while acknowledging the flexibility of the language itself - along with the practical consequences of each, were continually fine-tuned and adjusted over the 71 year project (the original estimate for creation of the OED was "no longer than 10 years"). For those of us looking to categorize and catalog our own site information, don't these balances sound familiar? As the LTER community moves with the tide of informatics to meet the present and future needs in the formidable realm of standards, we have the distinct advantage of facing the future armed with our past.