GEOTECHNICAL DATA DISSEMINATION IN THE NEESGRID

by

Hank Ratzesberger, Technical Staff Assistant

University of California at Santa Barbara

June 10, 2004

DISSEMINATION OF GEOTECHNICAL DATA ON THE NEESGRID

The NEESgrid is a collaboration of fifteen universities, of which three are providing equipment and conducting experiments involving geotechnical data. The University of California at UCLA, The University of California at Santa Barbara, and The University of Texas at Austin, in particular are conducting field site experiments in Soil-Foundation-Structure Interaction and Liquefaction. Each of these is conducting CPT, borehole and geophone recordings as well as seismometer/accelerometer recordings on soil and structures under natural and induced ground motion.

NEES equipment sites are run under the auspices of the NEES Consortium according to the mandate of the National Science Foundation. A key mandate is that all experiment data be thoroughly recorded and made publicly available to other researchers. Over the next ten years of operation, NEES should be a significant source of geotechnical and related data. Through its grid computing infrastructure, the NEESgrid will provide a single repository for all NEES experiments. However, the NEESgrid is still being developed and the technologies it uses are not all mature. Obtaining NEES data in any automated way may require that specialized tools be developed (or completed). Finally, there is no mandate for NEES to provide the data in formats that particular research communities have adopted, and so strategies and tools may yet need to be developed to provide in, for example, the COSMOS Geotechnical format.

Providing complete experiment data and metadata is a high priority of NEES and the NSF. The NEESgrid developers at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, recognizing that models for structural experiments would evolve, and wishing to capitalize on new

developments for the "Semantic Web" chose to record experiment metadata in the Web Ontology Language (OWL) format. This requires the development of an ontology for each experiment type. NEES has funded ontology development for centrifuges, tsunamis tanks and shake tables. No specific plan exists to develop an ontology for geotechnical, geophysical or soil-foundation-structure interaction.

The support within OWL for interacting with XML Schema documents, such as the COSMOS Geotechnical Schema is limited. Indeed, while it is a goal of OWL to co-exist with XML Schema "particles" or provide a basic translation from Schema to OWL, neither of these strategies currently exists.

This author was fortunate to participate in testing of the COSMOS Geotechnical schema and developed a simplified version of the schema that, it was hoped, could exist within the OWL Ontology. The World Wide Web Consortium is responsible for the development of OWL, and until specifications exist for the two XML syntaxes to cooperate, NEES will not be able to use the COSMOS Geotechnical format either as is, or modified specifically for this purpose.

Initially, it was assumed that the NEES sites would develop their own experiment metadata specifications. That task proved to be more involved that estimated, but NSF provided some extra resources, as mentioned. However, using the COSMOS Geotechnical schema as a model, it should be possible to create an OWL Ontology that is very similar to the schema. Ontology development is done using Protégé, which was developed at Stanford University.

A mandate of the NEESgrid is a repository of all experiment data and metadata. The San Diego Supercomputer Center (SDSC) will be the home of this repository during the proposed ten year life of the NEESgrid. All published data will be available through its NEES Point of Presence (NEESpop) on the

Internet. The NEESpop provides a web-based application that allows users to browse and query the repository and down load experiment data and metadata. Data will be available in standard file formats such as SEGD, mini-SEED and even as ASCII and Excel spreadsheets. Metadata, such as the experiment abstract, equipment specifications and configurations, applied filters and transformations will be available through the Metadata Browser in the NEESpop. Specific features and methods for searching metadata and acquiring data have not been developed or described as of this writing. Currently, only computer systems running the NEESpop software can access metadata or data.

The NEES data repository has been designed to become part of the future "Semantic Web" initiative and not to provide data in specific formats, even those created by recognized public authorities. Whether this decision is a shortcoming remains to be seen. The Semantic Web initiative could change the nature of the Internet in the same way the graphical browser did. In the mean time, providing a bridge from the OWL classes to the COSMOS Geotechnical Schema is, in the opinion of this author, a worthwhile project that interested researchers are invited to propose. For now, anyone can create a login on any NEESpop and, when the 3.0 version is released in July, examine the current ontology and help extend it to support geotechnical and geophysical experiment data.

REFERENCES

- NEES 2nd Annual Convention Meetings:
  http://www.nees.org/NCI/meetings/annual_mtg2/docs/proceedings2.pdf
- NEES Data Sharing and Archiving Committee:
  http://www.nees.org/NCI/committees/DSA/index.html
- Vision of NEES Collaboratory:
  http://www.nees.org/CD/task_groups/docs/VisionWhitePaperV3.pdf
- Web Ontology Guide:
  http://www.w3.org/TR/2004/REC-owl-guide-20040210/
- The Semantic Web:
  http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21
- Protégé:
  http://protege.stanford.edu
- San Diego Supercomputer Center
  http://www.sdsc.edu/