Data Management Plan

Data Description

The project will utilize the computational, experimental, and IT network facilities located at Lehigh University. Experimental, computational, and simulation data developed will be archived to enable re-use by researchers. The Lehigh EF web site will provide a local repository for experimental, computational, and simulation data for the project. Users will be able to upload data to the NHERI data repository through the internet connection between the Lehigh EF and NHERI.

The data management plan contained here is associated with physical experiments and numerical simulations performed by users of the facility. The researchers will be given training and assistance by the Lehigh EF IT Systems Manager, but will responsible for uploading all of their data, including: the metadata that describes the experimental setup report; unprocessed experimental data; converted experimental data; metadata that describes the numerical simulation models; numerical simulation model input data; numerical simulation model output data; numerical simulations processed data; pictures and video.

Data to be included under Data Management at the Lehigh EF includes the following categories:
- Unprocessed or raw experimental data in ASCII or binary format obtained by data acquisition and control hardware (Pacific Instruments, ServoTest, Wineman, Mathworks, National Instruments).
- Converted and derived data sets from unprocessed data obtained through computational software (MATLAB, Microsoft Excel).
- Computational models including analytical data in ASCII format.
- Experimental photos and videos through web cameras, point and shoot cameras and camcorders.
- Software developed which generates NHERI related project data.

All data will be made available raw and uncorrected, as well as in converted forms in engineering units. In addition, pictures and videos will be part of the data collection. Annotations will be used to indicate relevant information describing events during each test, and the data will be documented to enable a qualified researcher to reproduce the simulation or test and generate the derived data throughout the lifetime of data repository.

Data and Metadata Formats

The procedures for data archiving will follow the NEES Project Warehouse defined standards for metadata. The data hierarchy is defined below.

- *Project*: a Project represents an overall research project, which includes one or more experiments/numerical simulations.
    - *Experiment*: an experiment represents a physical test or a computational simulation.
        - *Trial*: a trial is associated with an experiment or a simulation. Multiple trials may be applied to the same experimental or simulation set up.
            - *Repetition*: a repetition is associated with a trial and represents the application of the same trial without any change to the test set up or trial parameters.

The format may change upon the request from the Cyberinfrastructure Awardee to update the Data Management plan.

Data Archiving on Local Repository

Lehigh University maintains a redundant and scalable network attached storage system (RTMDdrobo) which contains all data and documentation related to the EF. This system is tightly integrated into the IT

architecture and is available as a mapped system to any workstation or server. This system provides dual disk redundancy which ensures that the data remains safe and accessible to users even if as many as 2 drives fail at any given time. This system is protected by the EF firewall and the University firewall. Lehigh also utilizes an offsite backup system to guarantee that data is stored safely offsite in case of catastrophic disaster. Data is organized into a file-folder structure that mirrors the NEEScomm data model format and is uploaded to the central repository within 24 hours of test completion via a 1Gbps Internet2 connection or up to 10Gbps through the commodity Internet connection. Typically, files are named using a "testname_date.[ext]" format while folders are appropriately named for sets of trial data. Access to the physical RTMDdrobo system is limited to the Lehigh EF IT Manager and is handled with the utmost care and discretion for privacy, security and integrity reasons.

The Lehigh EF will comply with all requirements for post-award monitoring including annual reports providing progress on the data management and sharing of research products, publications and conference proceedings, and sharing and dissemination of results. Final project reports will be in compliance with NSF requirements, including the execution and any updating of the Data Management Plan.

The data acquired and preserved in the context of this proposal will be further governed by Lehigh University's policies on intellectual property, record retention, and data management. The details of this policy can be found at:
http://www.lehigh.edu/~policy/documents/records_retention_policy_06-24-2010_withappendix.pdf