

University of Florida NHERI Experimental Facility Data Management Plan

(Version 1.0, Revised 3/10/2017 by JAR)

The University of Florida (UF) NHERI EF is committed to ensuring that the project results, including experimental data, developed software, and hardware designs are properly archived and disseminated to the broader research community. The UF EF will work with its users (project PIs) to promote the use of the tools provided by the NHERI DesignSafe Cyberinfrastructure (CI) at the Texas Advanced Computing Center (TACC).

Data Types and Formats

- *Documentation* will be prepared to describe the tasks and outcomes for all facility objectives. These documents will consist of items such as instrumentation manuals, training materials, and classroom materials. Documents will be produced using MS Office products (Word, PowerPoint) and made available in write protected PDF format. Design documents will be prepared in CAD (e.g., SolidWorks), structural / finite element analysis (e.g., ADINA), and simulation software (e.g., Simulink)
- *Experimental data* will consist of both raw data for internal use (not for wide dissemination) and processed data in various forms including graphical, tabular, and software encoded. These data concern the development and calibration of the subsystems and model testing. Data will be stored using standard formats such as formatted text files (.txt), MS Excel (.xlsx), MATLAB (.mat), and LabView (.lvn). Large data sets will be stored as hierarchical data files (.hdf, .hdf4, .hdf5) or equivalent
- *Software* developed during this study may be produced to acquire and store data, control instrument's subsystems or to provide the integrated control and visualization platform. Software will be written in python, C++, MATLAB, and LabView. Software will not be published openly, but available upon request in an unprotected format.

Data and Metadata Standards

- *Metadata* describing the date/time of the experiment, sensor locations and specifications, model geometry and material, experimental conditions, and the research personnel involved will be stored.
- Project PIs will use the standard data models available on DesignSafe-CI developed for the UF EF resources.

Data Access and Sharing

- The EF will work with project PIs to coordinate and assure data storage and access as well as manage, store and disseminate results.
- The EF's copy of the data will be stored on hard drives located onsite at the Powell Laboratory
- The High Performance Computing Center (HPC) at the University of Florida (<http://www.hpc.ufl.edu/about/>) will be leveraged for storage and archiving of large data sets using a ZFS-based file system on enterprise-class SAS/SATA disk drives. All data will be backed up to this server, which will serve as the primary EF-located backup storage for the project.
- Data will be stored at UF for a minimum of three years after the conclusion of the award
- The EF will promote the use of DesignSafe-CI to store and publish data, document the project, provide technical specifications and metadata, and instruct interested parties on how to properly request and use the data.
- The use of DesignSafe-CI will facilitate scientists, engineers and other interested parties obtaining the data directory from the Data Depot within DesignSafe-CI. Intellectual property or first right to publication will be determined by the Project PIs and controlled through the PIs chosen data distribution method (e.g. DesignSafe).

- It is anticipated that dissemination of EF-collected data will be actively pursued by the project PIs in traditional academic forums such as conferences and workshops, and direct communication with peers within the research community (including those from non-Ph.D. and/or minority-serving institutions)

Data Re-use and Redistribution

- PIs can select one following licenses for their data distribution (offered as options by DesignSafe):
 - For datasets: ODC-PDDL and ODC-BY
 - For copyrightable materials (for example, documents, workflows, designs, etc.): CC0 and CC-BY
 - For code: any open, non-commercial license (for example, GPL)

Data Archiving and Preservation

- The NHERI CI will persistently maintain all data uploaded to Data Depot on storage resources at the Texas Advanced Computing Center, and these resources are redundant and geographically replicated