

NHERI Experimental Facility 2015-2019
University of California, San Diego (UCSD)

Data Management Plan

This Data Management Plan addresses the NSF policy that primary data “commonly accepted in the scientific community as necessary to validate research findings” is made available to the community at little or no cost to the PI or project. The end-to-end data management plan for the NHERI@UCSD EF will comply with UCSD standards of common practice (such as high standards for internal experimental reports and data archiving and UCSD Guidelines on Access and Management of Research Data [1]), as well as those set by the NHERI Cyberinfrastructure (CI) awardee for data and metadata protocols to facilitate archiving and automatic data uploading to the NHERI data repository for curation. The EF will ensure that all unprocessed data is reliably collected and is of the highest quality, and will work closely with the PIs and research team for each project, who will be ultimately responsible for NHERI data curation, to support their process of properly archiving and preserving their data. Expertise in Cybersecurity and Data Management is leveraged from the San Diego Supercomputer Center (SDSC) at UCSD, who provides cyberinfrastructure resources to scientists requiring massive computer and data-handling capabilities to conduct their research. The EF will interface extensively with the UCSD Administrative Computing & Telecommunications (ACT) group, who provide dependable, high quality, effective and efficient IT services and infrastructure to the UCSD community, and provide policies and standards. This will ensure compliance with UCSD and CI Awardee policies. The EF has experience with a similar strategy during its previous efforts as a facility within the NEES equipment network.

I. Types of Data

The majority of the data developed at the NHERI@UCSD EF will be experimental data collected through sensors and video. Data files are generated in binary format then converted to ASCII format to share with project PIs or their designees. Technical drawings showing the type and location of the sensors will be required from researchers for instrumentation of the specimen. Diagrams indicating location and angle of cameras or desired field of view will also be required for video. A list of channels will be generated from the data acquisition system indicating sensor type, calibration constants and other metadata required for archiving. In the case of hybrid simulation, numerical simulation data will also be included and stored on site by the PI. Achieved input loads and displacements from hybrid simulation tests will be provided by the EF. Additional data created by this project will be in the form of evaluation results of performance metrics, numerical analyses, surveys and responses, participant feedback and deliverables promised to the National Science Foundation (reports and reviews). These data will be captured as web site database files, images, spreadsheets, and text. The release of data regarding students or any human subjects is subject to policies and restrictions in protocols adopted by the relevant Institutional Research Board (IRB) and the Family Educational Rights & Privacy Act (FERPA) regulations.

II. Data and Metadata Standards

Based on past NEES data sharing and archiving policies, researchers will be supported in their requirement to provide all relevant information necessary to completely and accurately describe their experiments. This includes setup details such as testing methods, specimen drawings and sensor locations and specifications, photographs of the setup and instrumentation, as well as response data such as recorded and interpreted measurements, photographs post-test inspection and demolition, and with provenance, associated numerical simulations and models, video and other photos, and annotations, to ensure that they are available in a form that the engineering community can understand and readily use for current and future endeavors in research, education and practice. In addition to UCSD data standards of common practice [1], the EF will adhere to NHERI Data Sharing and Archiving policies and standards that are published to support metadata archiving and will provide necessary metadata to enable a qualified researcher to reproduce the test. This includes appropriate sensor sensitivity and filtering information, annotations describing events during testing such as the failure of a sensor or channel that may affect the data recorded.

Furthermore, documentation about the facility will be created using standard Microsoft Office applications (Word, Excel, and PowerPoint) and HTML formats to publish facility policies and procedures online. These documents will be shared in the latest software versions compatible to a variety of platforms (Mac OS, Windows, etc.). All applications and formats are widely used and easily exported for sharing. For the purpose of wider, long-term access, primary documents will be converted at regular intervals into pdf documents because they preserve the integrity of the original documents, and are easily

accessible by anyone with an Adobe Reader (free to download from the web). The metadata is mainly in the form of columns documented with headings and explanatory text.

Data Workflow The following steps will be implemented for each project carried out at NHERI@UCSD (note these steps follow preliminary meetings/teleconferences as necessary to develop proposals, and after assignment of test dates at the EF):

1. Research Team meets with the facility staff to discuss safety and testing plans (motion types, sequencing, amplitude, specimen variations and various phases in the testing, etc), and develop a preliminary plan for sensor installation. Based on this, a detailed instrumentation plan is developed including drawings with sensor and video locations.
2. Except for strain gauges and special sensor, NHERI@UCSD staff will install all sensors. Sensors will be installed based on drawings provide by the research team, which will ensure completeness.
3. All sensors will be connected to the DAQ by the site staff, who will also enter all necessary sensor metadata into the DAQ support computer. The sensors will be checked and malfunctioning sensors or channels will be fixed or replaced.
4. Staff will train the designated research team member(s) on data file format and data transfer will occur by providing metadata files generated by the DAQ including a sensor list with calibration constants and units.
5. For LHPOST, prior to a shake table test, which is anticipated to cause inelastic response of the specimen, preliminary low-level tests will be conducted and sensor data will be collected and checked to verify each sensor.
6. After each test, unprocessed ASCII data will be immediately given to the research team and archived in a local storage system; after each test day, data will be uploaded to an off-site storage and NHERI data repository; both of which will be immediately accessible to the research team.

III. Policies for access and sharing and provisions for appropriate protection/privacy

Researchers will follow all UCSD and NHERI specified data policies regarding access and sharing. Typically within 6-12 months from the end of an experiment or hybrid simulation, researchers will be required to properly evaluate the quality of data, provide documentation for curation, and submit corrected and derived data to the NHERI Data Repository. These data are not publicly available until either the researcher chooses to make the data public or the curation process is complete.

The UCSD Human Research Protections Program (IRB) will clear any shared data that involves human subjects from an ethical and privacy perspective. The shared data is not 'personal data' in terms of the Data Protection Act 1998 (the DPA) or equivalent HIPAA requirement). The data are neither copyrighted nor licensed. These data, as well as all the documents relating to the project, including publications, reports, papers, and project summaries, will be captured on systems at NHERI@UCSD and provided as resources to the NCO and CI for inclusion on the NHERI website. In addition, the center will also maintain and update supporting documentation related to all facility specifications, proposal preparation instructions, and system maintenance and support policies.

IV. Policies and provisions for re-use, re-distribution

All materials will contain acknowledgement of NSF support as per NSF policy: "*This material is based upon work supported by the National Science Foundation under Grant No. (NSF grant number).*" Additionally, materials will contain a disclaimer that "*any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.*" In keeping with standard ethical practices, it is expected that subsequent users of the data will acknowledge the source.

V. Plans for Archiving and Preservation of access

The facility has a scalable data storage system that contains all test data produced at the facility. This system is protected by firewalls. Experimental data (e.g. measurements, video, and photographs, field notes) will be stored in a local repository as they are recorded. In the case of hybrid simulations, computational data associated with the simulation will also be stored in an initial repository as they are generated. The facility also utilizes an offsite backup provider to guarantee that data are stored safely in case of a catastrophic disaster. The data are uploaded to the off-site storage system at the end of each test day, and will be uploaded to the NHERI Central Data Repository within one week of test completion by the table operator. At the end of each test day, data are transferred to the researchers through a portable storage device.

[1] <http://blink.ucsd.edu/research/policies-compliance-ethics/guidelines.html>