



This document describes policies and guidelines of the George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES) program for archiving and sharing data resulting from NEESR funded projects and those projects funded by other agencies under the NEES Shared-Use program, and for other relevant data to be shared through the NEES Data Repository.

Policies for Data Archiving and Sharing

1. Introduction

1.1 Background and Purpose

NEEScomm (NEES Community and Communications) facilitates the use of the NEES Equipment Sites and establishes data archiving and sharing policies for all data submitted to and archived in the NEES Data Repository. The NEES Data Repository is also known as the Project Warehouse. The building of this resource requires a partnership between the researchers conducting the projects, the Equipment Sites providing technical and logistical support and NEEScomm as administrator for the Project Warehouse and NEEShub. By participating in NEES repository, each of these partners accepts certain responsibilities that ensure establishment of this valuable resource for use by the international earthquake engineering community.

The overall purpose of this document is to set forth policies that encourage, ensure and facilitate full and open access to high quality data archived in the NEES Data Repository.

1.2 Roles and Responsibilities

NEEScomm is committed to promoting the sharing and reuse of high quality data generated within the scope of the NEES mission. A core responsibility of NEEScomm includes establishing, maintaining, and preserving well-documented, properly safeguarded data archives and developing standards for data access by members of the scientific and engineering research communities. It is the role of the sites to ensure that the data populating the NEES Data Repository is responsibly obtained. It is the role of the researchers to ensure that the data populating the NEES Data Repository is complete in its documentation for current and future users.

1.3 Data Goals of NEES

1.3.1 Data and documentation will be archived and preserved

Each project involving experimental simulations or the files documenting actual earthquake events is a valuable resource for the earthquake engineering community. It is of utmost importance that all data is reliably collected, archived and preserved. Researchers must provide all relevant information to provide a complete and accurate description of an experiment, including setup details such as testing methods, specimen drawings and sensor locations and specifications, as well as response data such as recorded and interpreted measurements and with provenance, associated numerical simulations and models, video and photos, and annotations, to ensure that they are available in a form that the earthquake engineering community can understand and use for current and future endeavors in research, education and practice. Researchers are to upload this information in a timely manner; NEEScomm will ensure that reliable and efficient mechanisms are in place to do so. Backup and recovery plans will be established and followed at key data collection points.

1.3.2 Archived data will be of high quality

A NEES Data Repository populated with high quality data from research teams will be a significant resource for the earthquake engineering community, facilitating the mitigation of risks associated with earthquakes around the world. Thus, the data submitted for curation must be responsibly generated and must be well



documented. Sufficient information must be provided about the data and their creation to ensure that any qualified user could reproduce the experiment and/or simulation.

1.3.3 Archived data will be shareable and searchable

Ultimately the goal is to share the data placed in the NEES Data Repository with the broader earthquake engineering community along with educators, practitioners and the public. However, we recognize that the researchers who create and collect the data should benefit first from that data (for a finite period of time) before the data is released to the broader community.

1.3.4 Archived data will be preserved

Upon uploading of the data, documentation, and metadata by the research teams, NEEScomm is committed to the long-term preservation of data and digital objects in the NEES Data Repository. NEEScomm will ensure that files associated with research projects in the NEES Data Repository will remain accessible and readable over time.

1.3.4 Establishing a trustworthy repository

NEEScomm will follow the OAIS (Open Archival Information System) reference model (<http://public.ccsds.org/publications/archive/650x0b1.pdf>) to create conditions for a trustworthy digital repository of research data.

2. Policies for Data Archiving and Sharing

This section describes the policies for the archiving and sharing of data that is to reside in the NEES Data Repository.

2.1 Applicability

These policies apply to all information archived in the NEEScomm Data Repository. These include data resulting from NEESR funded projects and those projects funded by other agencies under the NEES Shared-Use program, and for other relevant data related to the scope of NEES to be shared through the NEEShub.

Prior to the start of work can begin at a NEES Equipment Site, an Equipment Site Utilization Form (ESUF) must be completed by the PI, approved by the NEES Equipment Site, and submitted to NEEScomm. The ESUF contains three data-specific sections. These sections are:

- 1) **Information Technology and Data Archiving Roles and Responsibilities**
- 2) **Expected Data Volumes**
- 3) **Exceptions to NEES Data Sharing and Archiving Policies and Guidelines (as needed)**

2.2 Quality Assurance

All researchers and NEES Equipment Site personnel are responsible for providing high quality data. The quality of the information is dependent on the use of appropriate testing procedures in the generation of the data (e.g., sensor sensitivity, filtering used, etc) as well as the completeness of documentation provided about the tests performed. Annotations should be used to indicate relevant information describing events during the test (e.g. in the case of a sensor failure during a test). The project data must be documented to enable a qualified researcher to reproduce the test, and also to reproduce generation of the derived data submitted throughout the lifetime of the NEES Data Repository.

Tools and training will be available through NEEScomm to facilitate documentation.

2.3 Archiving

This section describes the data flow from production through packaging to curation and to sharing with the broader community through the NEES Data Repository.

Experimental data (e.g. measurements, video, and photographs) must be stored in a local repository **as they are** recorded. In the case of hybrid simulations, computational data associated with the simulation should also be stored in an initial repository **as they are** generated. Researchers are encouraged to upload unprocessed data to the Project Warehouse immediately after a test, where that data can readily be made available to the research team.

NEES Equipment Sites are to provide and maintain offsite backup and storage of all data acquired at their Site. Unprocessed data, taken directly from an instrument or data acquisition system, must be stored offsite within 24 hours after each experiment unless otherwise justified in the Exceptions section of the ESUF. Sites are encouraged to use the Project Warehouse itself for offsite storage. Each NEES Equipment Site shall keep a current plan on file with NEEScomm detailing how the Site is implementing the initial repository and the offsite data storage requirement.

Researchers have six months from the end of the experiment or hybrid simulation, to properly evaluate the quality of the data, provide documentation for curation, and submit corrected and derived data to the NEES Data Repository. This data is not publicly available until either the researcher chooses to make the data public or the curation process discussed subsequently is complete. Where a series of tests are conducted, the six month deadline for all tests is from the last test in the series.

Researchers are further encouraged to include additional interpretations and analyses that will help users understand and interact with the measured data.

Data submitted to the NEES Data Repository will be curated by NEEScomm within 3 months of the date that it is fully checked and documented by the researchers, who may be contacted by the NEES Data Curator if additional information is needed. Responsiveness to the Data Curator's inquiries is important to expedite curation.

No later than 3 months after the termination date of the grant, researchers should provide an Executive Summary for the project, which describes the data obtained and its relevance to future users of that data.

2.4 Accessibility and Sharing

NEEScomm expects researchers using the NEES Equipment Sites to make their research data available in a timely and responsible manner to the scientific community for subsequent research with as few restrictions as possible in keeping with the following:

1. Data will be made accessible to project participants as it is uploaded. Project PIs will be able to provide particular users access to the data before it is publicly accessible. System Administrators and Data Curators will be given read-only access to the data and metadata before it is made public. Strict confidentiality about the data will be maintained.
2. All data modifications, including those for correcting errors and for curation, must be approved by the project PIs or other person responsible for the project. Once approved, such modifications become part of the NEES Data Repository.
3. No later than 12 months after the end of an experiment, information placed in the NEES Data Repository archive and curated will be automatically released to the broader community.
4. The automatic release period may be extended under exceptional circumstances. The request for exceptions should be submitted as part of the Exceptions section of the ESUF, which is subject to the approval of the NEES Equipment Site and NEEScomm.
5. In the event proprietary information must be used for the conduct of a project, researcher shall provide justification for the use of said information in the Exceptions section of the ESUF. The NEES Equipment Site and NEEScomm shall have the right to deny the request.

2.5 Attribution

Researchers are responsible for ensuring that all data files are internally documented to include appropriate information about the NEES Equipment Site, the research project and sponsoring agency information in documentation.

2.6 Acknowledgements

All data stored in the NEES Data Repository should be appropriately acknowledged. Credit should be given to the researchers, equipment site(s), and sponsoring agency. Every project made available in the NEES repository contains suggested citation.

2.7 Compliance

The milestones and data-specific information in the ESUF will be monitored by NEEScomm. Failure to comply with the approved ESUF may disqualify a researcher from future use of NEES Equipment Sites.

2.8 Data Curator contact information

Questions about Data Curation may be directed to the NEEScomm Data Curator, Stanislav Pejša at spejsa@purdue.edu.