# University at Buffalo's NEES Equipment Site

# Data Management

## Jason P. Hanley
### IT Services Manager

**Structural Engineering and Earthquake Simulation Laboratory,**
Department of Civil, Structural and Environmental Engineering

# Overview

- Data Goals
- Data Flow
- Data Organization
- Data Structuring

# Data Goals

- Collect and preserve all data necessary to reproduce an experiment - No data lost

- Use the data:
  - Visualize
  - Analyze
  - Search
  - Compare

- Share data with:
  - Researchers
  - Students
  - Industry
  - Public

# **Types of Data**

- Data is the set of results produced by a numeric simulation or the readings from sensors in a physical test.

- To fully describe what this data means, more information is required, such as
  - Simulation or physical loading input files
  - Specimen description, figures, and photos
  - Instrumentation layout and calibration
  - Description of test procedures

- This is called **Metadata** and refers to the *Data about Data* collected during an experiment.

# Data Terminology

- Unprocessed: Raw data directly from the data acquisition systems
  - Sensor data
  - Audio
  - Video
  - Images
- Converted: Data converted to standard file formats and in engineering units
- Corrected: Data that has been revised to compensate for errors
- Derived: Data that has been further processed to plot, compare, and analyze results

# Data and Metadata Formats

- Word processing files
  - Abstract, Project description & design
- Presentation files
  - Project proposal & description
- CAD files
  - Structural drawings
- Text files
  - Input motion, data, calibrations
- Spreadsheet files
  - Data, graphs, calibrations, test schedules
- Image files
  - Specimen & test set-up pictures
- Video files
  - Video observations
  - Video data
- Simulation files
  - Input files
  - Output data

# Data Flow

- **Prepare Experiment**
- **Acquire Data**
- **Organize Data**
- **Structure Data**
- **Curate Data**
- **Publish Data**
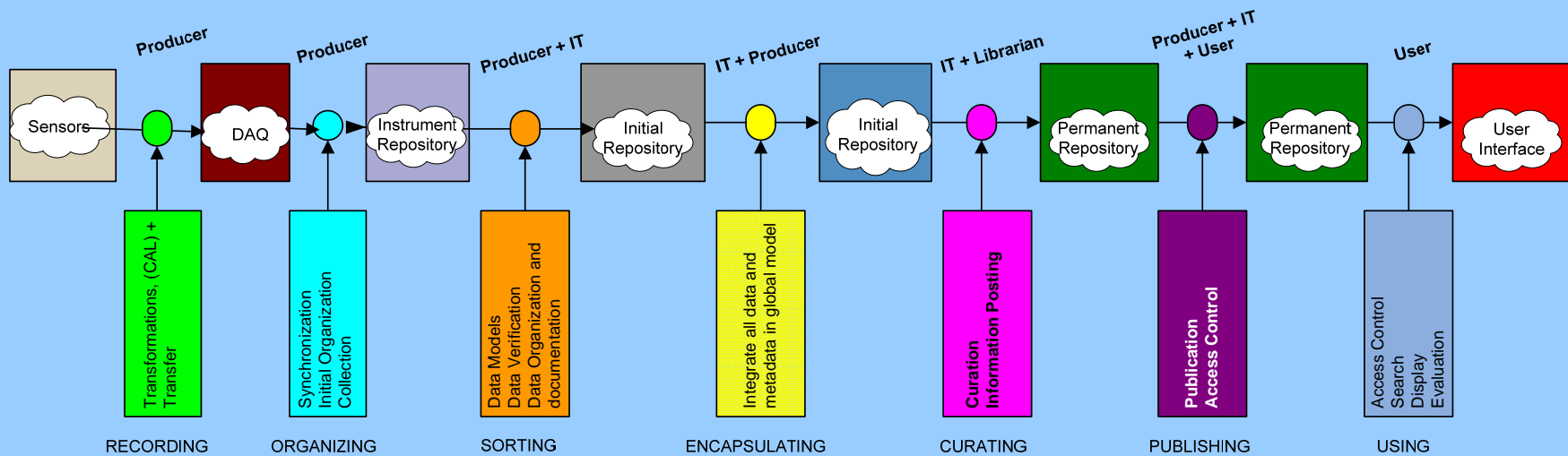
# Prepare Experiment

- Document information about an experiment
  - Specimen drawings and specifications
  - Instrumentation plan
  - Data acquisition systems
  - Loading systems
  - Test plan

# Acquire Data

- Sample data from sensors during a test
- Record data to the instrument repository
  - The instrument repository is local storage on the data acquisition and control computers
  - There can be more than one instrument repository
- Record metadata from data acquisition and control systems
  - Sensors list, channel list, calibrations, other configuration
  - Loading history
- Data and metadata will be in the native (unprocessed) formats used by the data systems
  - Note any processing done automatically by the data system

- In most cases, these tasks are handled automatically or by the lab staff, but you still share the responsibility to make sure they are done correctly

# Organize Data

- Collect all the data and metadata from the test and store it in the local repository
- Convert all data to engineering units
- Converts all data and metadata to standard formats

- Organize data and metadata in a hierarchy for easy mapping to the NEES data model

- Work with an IT specialist for this task

# Local Repository

- Located here at SEESL

- Where all data and metadata from a test is stored

- 3.5 TB of Network Attached Storage (NAS)
  - Backed up daily

- Accessible over the network
  - Map directly to your local file system
  - Browse on our website

# Data Hierarchy
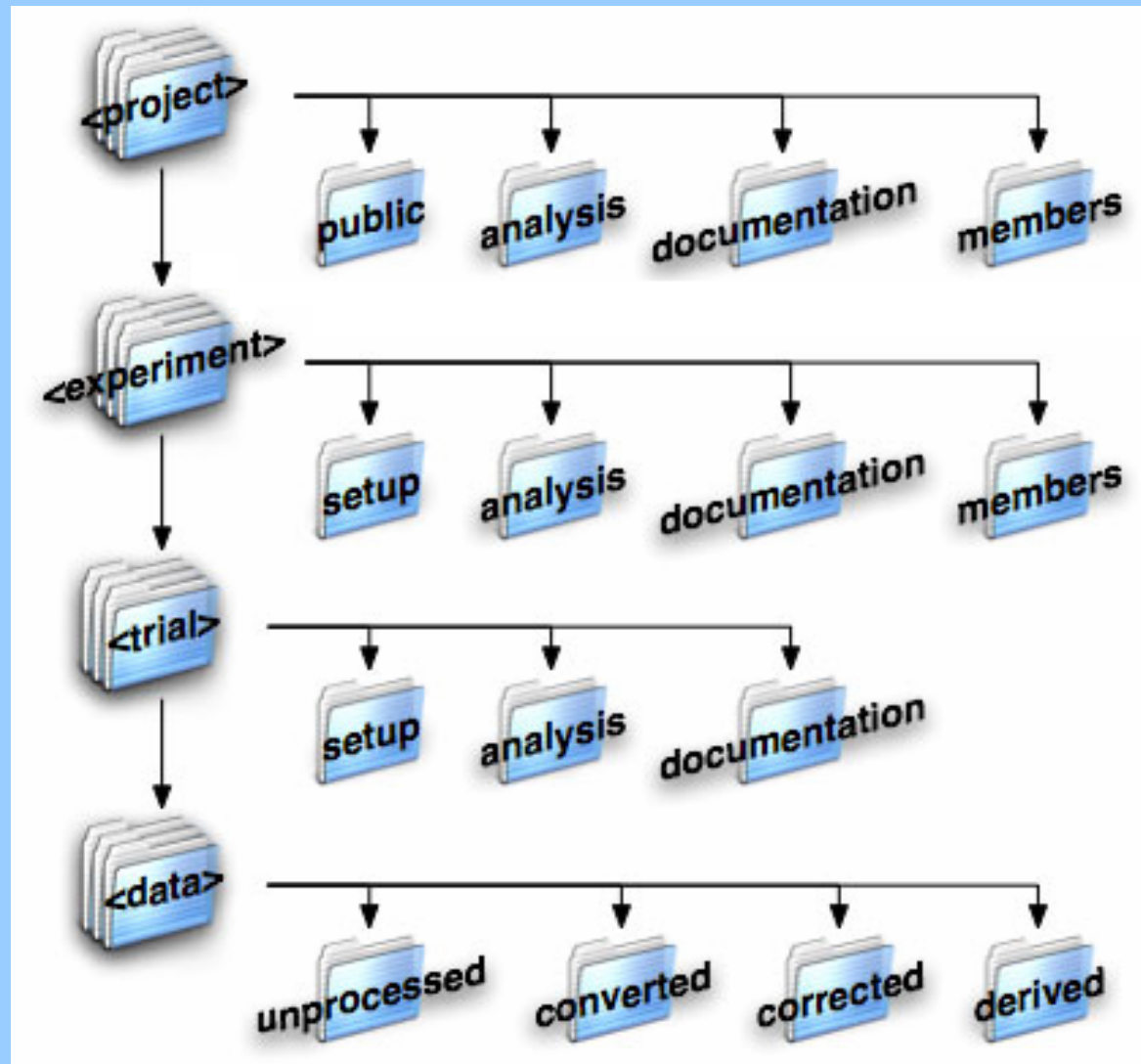
# Data Hierarchy (cont...)

## Project

- Name, nickname, objective, description, funding organization

- Members

- Documentation

  – Proposal, reports, papers, posters

  – Collaboration – presentations, emails, meeting minutes, notes, conversations, chats

- Experiment plan – name, date

# Data Hierarchy (cont…)

Experiment

- Name, objective, description, date/time, organizations, facilities
- Members
- Setup
  - Model
    - Structural drawings and specifications
    - Material and component properties
    - Scale factors
  - Instrumentation
    - Sensor location plan – name, type, location, orientation (list and drawing)
  - Data acquisition
    - Equipment
    - Channel setup
  - Loading system
    - Equipment
    - Input motions
- Documentation
  - Implementation – notebook, log of experiment
- Trial plan – name, date, time

# Data Hierarchy (cont…)

<u>Trial</u>

- Name, objective, description, date/time

- Setup

  - Channel list – sensor location, sensor, conditioning, daq channel, calibration

  - Input motion

- Data - unprocessed data, converted data, corrected data, derived data

- Documentation

  - Data processing performed

# Structure Data

- Map metadata organized in hierarchy to data model

- Upload data and metadata to the permanent repository

- Ensure all software used to produce data documented and is available


- Work with an IT specialist for this task

# A Data Model for Metadata

- Metadata needs to represented in some way.

- A **Data Model** is the specification of the format used to represent the metadata. It gives a structure to this metadata and creates relationships between different pieces of metadata.

- It is a standardized format so metadata created by others can be searched, viewed, and reused in the same way.

- A data model allows tools to be written to this standard that allow users to visualize and analyze this metadata is without worrying about which format it is in.

# NEES Data Model

- **Project:** Entity for managing data and metadata for a research project or other activity

- **Experiment:** Data and metadata necessary to describe the specimen, instrumentation, data acquisition, and loading system setup

- **Trial:** Data and metadata for a specific test
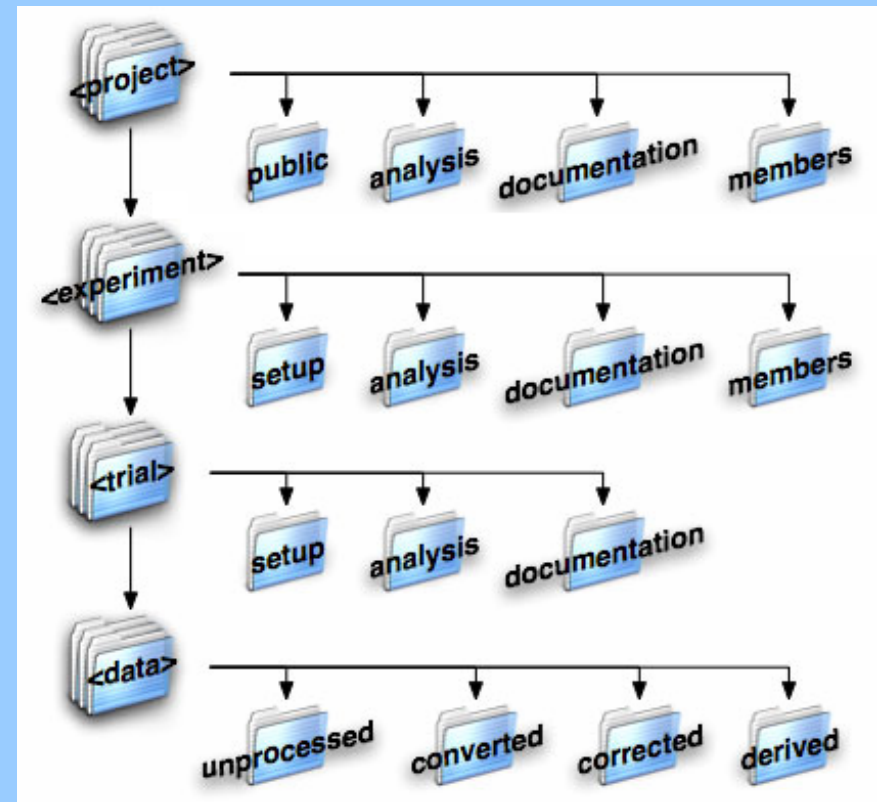
# Permanent Repository

- Long lasting data repository
- NEEScentral – Managed by NEESit
  http://central.nees.org/
- A representation of the data model
  - An interface for browsing and uploading data and metadata

# How to use the Data Model

- Pieces of data organized using the given hierarchical structure must be mapped into their respective metadata elements in the data model.

- These mappings allow for the metadata to comply with the standard set by the data model and provide an easy way to enter data into the model.

- For example, each experimental test would map directly to a trial in the data model.

# Curate Data

- Submit structured project to data curator
- Curator assess the projects conformance to the data model and data policies
- Work with curator to ensure a certain level of conformance, revise based on feedback
- Curator assigns a conformance level

# Publish Data

- Release the curated data set (project) to the public

# Additional Resources

- Managing Research Data
  http://nees.buffalo.edu/training/data/

- NEEScentral User's Guide
  http://it.nees.org/library/data/neescentral-users-guide.php

- NEES Data Sharing and Archiving Policies Guidelines
  http://www.nees.org/Governance/Policies/20050511_NEESinc_DSAPG.pdf

# Thank You !

# Questions ?