George E. Brown, Jr. Network for Earthquake Engineering Simulation

Publication of Research Data in the NEEShub

Stanislav Pejša, NEEScomm Data Curator

HUBbub 2013 Indianapolis, IN 2013-09-06

















Why Publish Data?

- Sharing data can be a source of recognition
- Incentivizing affect on promotion
- A tool for research assessment
- Can increase the citation rate
- Other possible use of resources
- Fosters responsible scholarship
- Strengthening open science
 - Global access
 - Protection against fraud
- Efficiency in use of scientific resources
- Enables new discoveries multiple perspectives

Costas, R., Meijer, I., Zahedi, Z. and Wouters, P. (2013). The Value of Research Data - Metrics for datasets from a cultural and technical point of view. A Knowledge Exchange Report, available from www.knowledge-exchange.info/datametrics



But Also ...

The White House wants you to do it...

The Administration is committed to ensuring that, ... the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

OSTP (February 22, 2013): Memorandum for the Heads of the Executive Departments and Agencies. http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf





Because ...

The White House wants you to do it... Because...

Policies that mobilize these publications and data for re-use through preservation and broader public access also maximize the impact and accountability of the Federal research investment. These policies will accelerate scientific breakthroughs and innovation, promote entrepreneurship, and enhance economic growth and job creation.

OSTP (February 22, 2013): Memorandum for the Heads of the Executive Departments and Agencies. http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf





...And Is Doing It

Executive Order -- Making Open and Machine Readable the New Default for Government Information

To promote continued job growth, Government efficiency, and the social good that can be gained from opening Government data to the public, the default state of new and modernized Government information resources shall be open and machine readable. Government information shall be managed as an asset throughout its life cycle to promote interoperability and openness, and, wherever possible and legally permissible, to ensure that data are released to the public in ways that make the data easy to find, accessible, and usable.

The White House. Office of the Press Secretary (May 09, 2013). Executive Order -- Making Open and Machine Readable the New Default for Government Information. http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp-public access memo 2013.pdf





G8 also Wants You to Do It

..

- 7) We, the G8, agree that open data are an untapped resource with huge potential to encourage the building of stronger, more interconnected societies that better meet the needs of our citizens and allow innovation and prosperity to flourish.
- 8) We therefore agree to follow a set of principles that will be the foundation for access to, and the release and re-use of, data made available by G8 governments. They are:
 - Open Data by Default
 - Quality and Quantity
 - Useable by All
 - Releasing Data for Improved Governance
- Releasing Data for Innovation
 G8 Open Data Charter and Technical Annex (2013, 18 June).

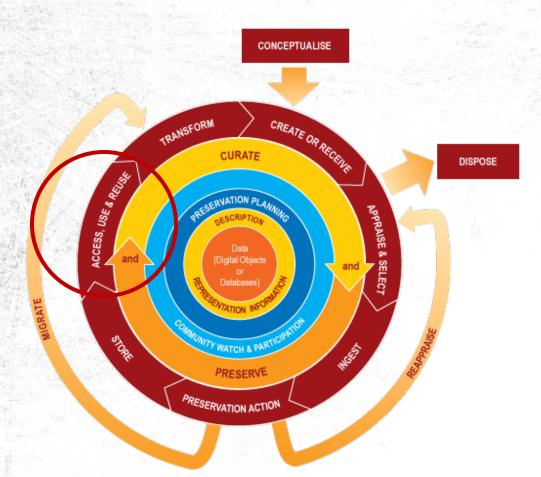
https://www.gov.uk/government/publications/ open-data-charter/g8-open-data-charter-andtechnical-annex







DCC Curation Life-Cycle Model



http://www.dcc.ac.uk/resources/curation-lifecycle-model

Data

Digital objects

Full Lifecycle Actions

- Description and Representation information
- Preservation planning
- Community watch and participation
- Curate and Preserve

Sequential actions

- Conceptualise
- Create or receive
- Appraise and select
- Ingest
- Preservation action
- Store
- Access, use, re-use
- Transform

Occasional actions

- Dispose
- Reappraise
- Migrate



What Is NEES?

14 engineering laboratories Shake Tables

- University at Buffalo
- UC San Diego
- UN, Reno

Tsunami Wave Basin

OSU

Geotechnical Centrifuges

- RPI
- UC Davis

Field Experiments

- UC Los Angeles
- UC Santa Barbara
- UT at Austin

Large Scale Laboratories

- Cornell University
- Lehigh University
- UC Berkeley
- UIUC
- UM, Twin City

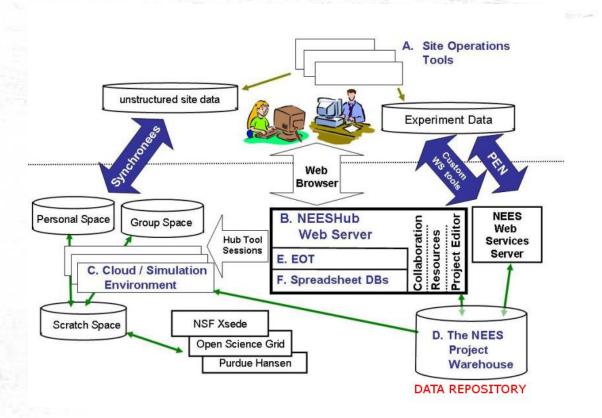




Cyberinfrastructure at NEES

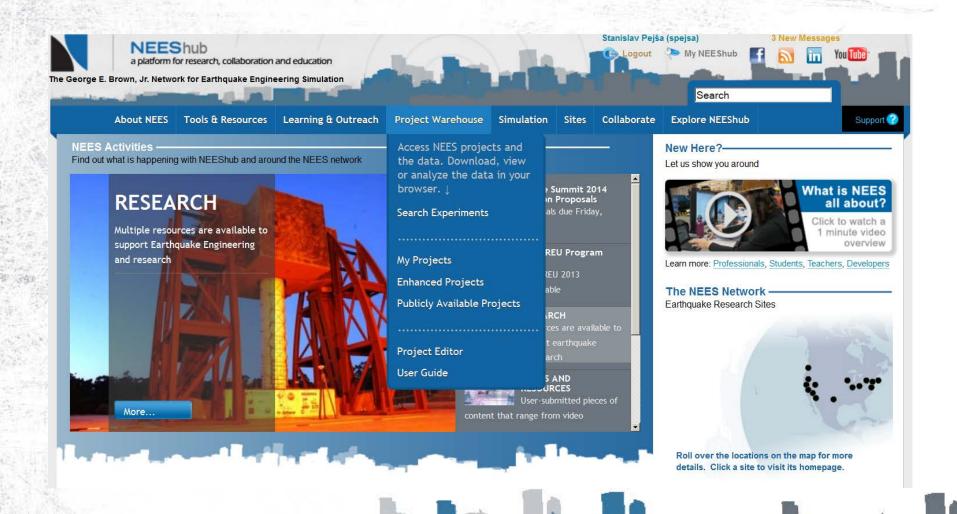
NEES cyberinfrastructure:

- A. Site Operations Tools
- B. The NEEShub Web Server
- C. Cloud / Simulation Environment
- D. The ProjectWarehouse NEESData Repository
- E. Education, Outreach, and Training (EOT)





NEEShub - www.nees.org



NEES

Data Archiving at NEES

Who

research team, site personnel, curator, NEEScomm

What

- sensor measurements
- sensor calibrations
- observations
- analyses
- numerical simulations
- reports (including publications and presentations)

When

- Dates are stated in the Data Sharing and Archiving Policies (1 month, 6 moths, 12 months)
- For as long as the data are useful ~ indefinitely ~ for 20 years

Where

Project Warehouse http://nees.org/warehouse/welcome

Why

- increases researcher's impact
- saves work, time, money
- facilitates knowledge transfer
- maintained authenticity and integrity of data
- good practice
- advances research

NEES

What kind of data?

- diverse
 - shared facilities, not always practices
 - research domain
 - structural engineering
 - geotechnical engineering
 - geophysical research
 - material engineering
 - tsunami research
 - type of data
 - experimental
 - observational
 - computational
- increasingly complex
 - number of sensors
 - interdisciplinarity
 - experimenting with computational modeling



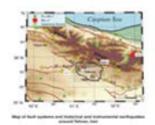
Data Publication in NEEShub

- All recently curated experiments
 - Have assigned DOI
 - Have improved metadata that facilitate discovery
- Datasets are considered published information products and NSF now allows listing information products in researchers' bio sketches.
 "Acceptable products must be citable and accessible including but not limited to publications, data sets, software, patents, and copyright"

http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg 2.jsp#IIC2fic

- The Earthquake Spectra journal is accepting a new type of manuscript called Data Papers.
 - Peer-reviewed papers that describe datasets of interest to the earthquake community
 - Data must be publically available with a Digit Object Identifier (DOI)
 - Submit soon for the inaugural issue of Data Papers http://earthquakespectra.org/page/data_papers







Citation and Attribution

Recommended citation format

Researcher 1, Researcher 2, Researcher 3 (YYYY), "Experiment Title" Network for Earthquake Engineering Simulation (distributor), Dataset, DOI:10.4231/D3SQ8QH1F

 Users of the data are expected to cite the data sets they used in the recommended format as shown above and also include an acknowledgement to the NEES Data Repository.

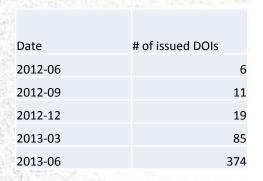
To acknowledge the NEEShub Data Repository:

The facilities of the George E. Brown Network for Earthquake Engineering Simulation (NEES) Data Repository were used for access to data and metadata used in this study (https://nees.org/warehouse/welcome). The NEES Data Repository is funded through the National Science Foundation and specifically the CMMI Directorate through the National Science Foundation under Cooperative Agreement Number CMMI-0927178



DOI – Digital Object Identifier

Number of issued DOIs since 2012-06-30





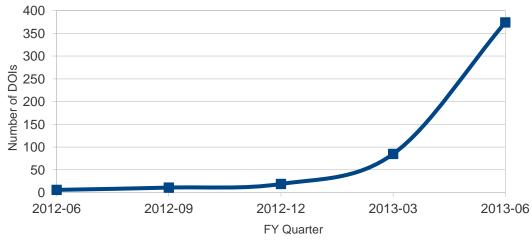
NUREG/CR-XXXX CCEER 13-09

Large Scale Earthquake Simulation of a Hybrid Lead Rubber Isolation System Designed with Consideration of Nuclear Seismicity

Manuscript Completed: October 3, 2012 Date Published: May, 2013

Prepared by K. L. Ryan, C. B. Coria, N. D. Dao

Center for Civil Engineering Earthquake Research (CCEER) University of Nevada, Reno/MS 0258 Reno, Nevada 89557-0258



- Ryan K, Sato E, Sasaki T, Okazaki T, Guzman T, Dao N, Soroushian S, Coria C (2013a).
 "Full Scale 5-story Building with Triple Pendulum Bearings at E-Defense", Network for Earthquake Engineering Simulation (database), Dataset, DOI:10.4231/D3X34MR7R
- Ryan K, Sato E, Sasaki T, Okazaki T, Guzman T, Dao N, Soroushian S, Coria C (2013b).
 "Full Scale 5-story Building with LRB/CLB Isolation System at E-Defense", Network for Earthquake Engineering Simulation (database), Dataset, DOI:10.4231/D3SB3WZ43
- Ryan K, Sato E, Sasaki T, Okazaki T, Guzman T, Dao N, Soroushian S, Coria C (2013c).
 "Full Scale 5-story Building in Fixed-Base Condition at E-Defense", Network for Earthquake Engineering Simulation (database), Dataset, DOI:10.4231/D3NP1WI3P



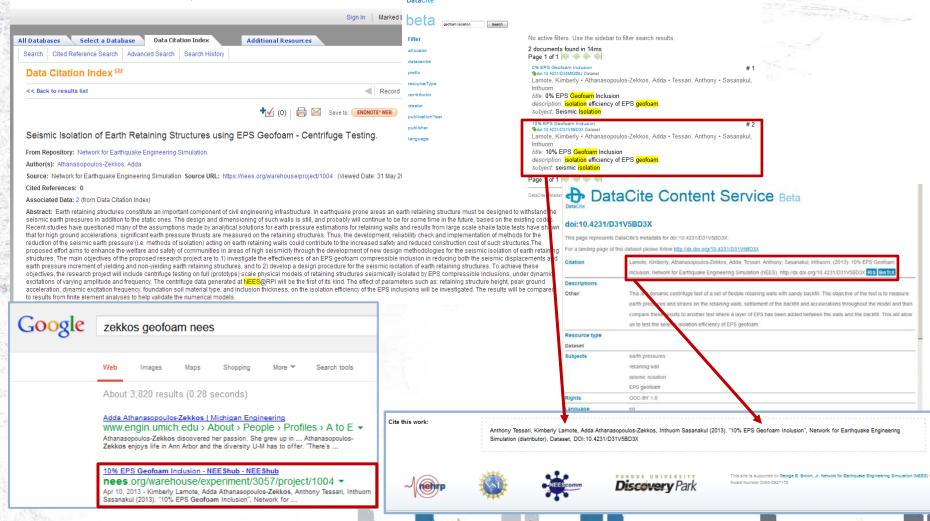
Exposure of EE Research Data

WEB OF KNOWLEDGE™

DISCOVERY STARTS HER



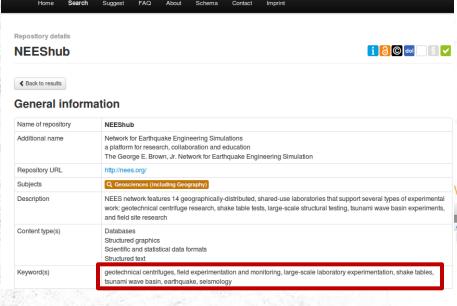
Options | Advanced Search | About Us | Contact | Hel





Repository for EE Data









Description

The George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES) aims to assis in the development of innovations related to construction and design that minimizes or prevents damage during natural disasters such as earthquakes or tsunamis. The NEES provides access to 14 shared-use laboratories that support experimental work in such areas as geotechnical centrifuge research, shake table tests, large-scale structural testing, tsunami wave basin experiments, and field site research.

Access: Open

Start Date: 2004

WEB OF KNOWLEDGE™

DISCOVERY STARTS HERE



Network for Earthquake Engineering Simulation.

Group Author(s): Network for Earthquake Engineering Simullation

Source: Network for Earthquake Engineering Simulation

Source URL: http://nees.org/ (Viewed Date: 22 Nov 2012) Published Year: 2004

Cited References: 0

Associated Data: 12 (from Data Citation Index)

Abstract: The Network for E-arthquake Engineering Simulation flourises on accelerating innovations in infrastructure design and construction practices for minimize damage during earthquakes on proceedings of the certainties data for expectant of the state of the certainties data for expectant of an animanical studies. The data in the Project Warehouse is the certainties data from experimental and unconstruct studies. The data in the Project Warehouse is determined and provided and the project Warehouse is determined and unconstruct studies. The data in the Project Warehouse are associated with research projects funded by a variety of agencies, including the National Science Foundation (NSF), and include experiments performed at NEES and non-NEES equipment after 2000 NEESHouse.

Document Type: Repository

Accession Number: DRCLDATA2013064003195872

Language: English

Addresses

Purdue University Discovery Park, 207 S. Martin Jischke Drive, West Lafayette, IN, 47907, USA

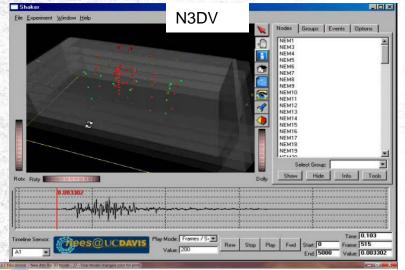
E-mail Addresses: neescomm@purdue.edu

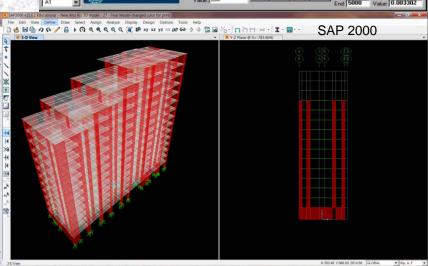
Web of Science Category: Construction & Building Technology, Architecture; Engineering, Multidisciplinary

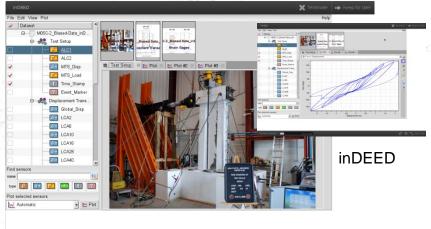
ubject Area: Construction & Building Technology; Architecture; Engineering

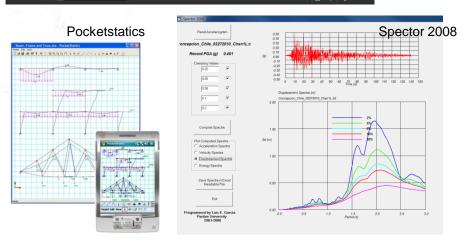


Visualisation



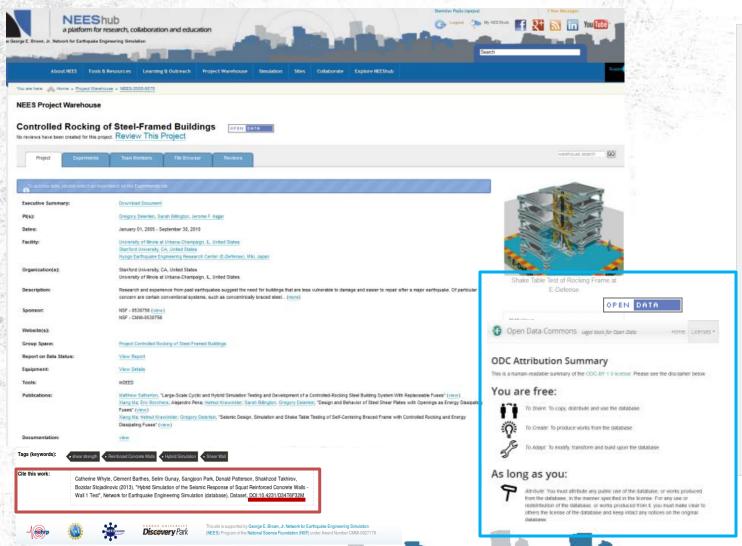








Curated Experiment

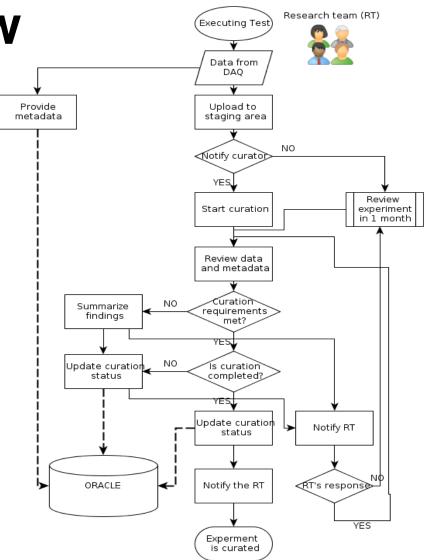






Curation Workflow

- curation is a process
 - starts early
 - "exit" interview
 - data upload
 - reminders
 - data review
 - experiment review
 - copyright compliance
 - preservation
 - DOI assignment



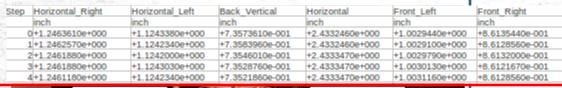


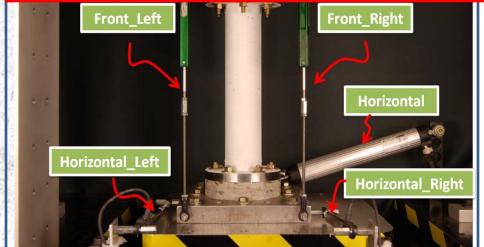
Curation as Quality Assurance

- Content
 - dependent on human monitoring
 - Metadata
 - Completeness
 - Based on standards and requirements
 - » On the level of research hierarchy
 - » Timeline
- Technical
 - machine-actionable
 - Formats
 - Interoperability
 - Accessibility
 - Preservability
 - File integrity



Understandable data





Hybrid Simulation LVDT Set-up

▼ Hybrid Simulation LVDT Set-up

	Channel Name	Label	Sensor Type	Comment	Orientation	XYZ Coordinates
	Horizontal_Right	Horizontal_Right	LVDT		0,1,0	-7.21 in, 12.112 in, 1.65 in
	Horizontal_Left	Horizontal_Left	LVDT		0,1,0	6.709 in, 12.014 in, 1.717 in
	Back_Vertical	Back_Vertical	LVDT		0,0,1	1 in, -4.65 in, 19.875 in
	Horizontal	Horizontal	LVDT		1,0,0	-14.913 in, -8.579 in, 5.084 in
	Front_Left	Front_Left	LVDT		0,0,1	4 in, 4.15 in, 19.875 in
	Front_Right	Front_Right	LVDT		0,0,1	-2 in, 4.15 in, 19.875 in
- 1						

Metadata need to be:

- meaningful
- purposeful
- consistent
- accurate
- predictable
- "standardized"

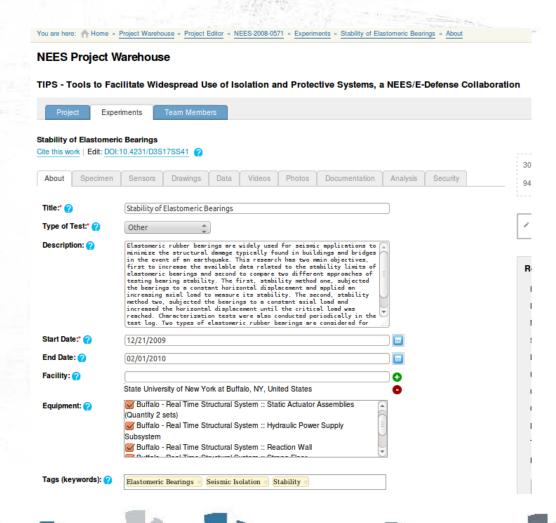
Relationship among:

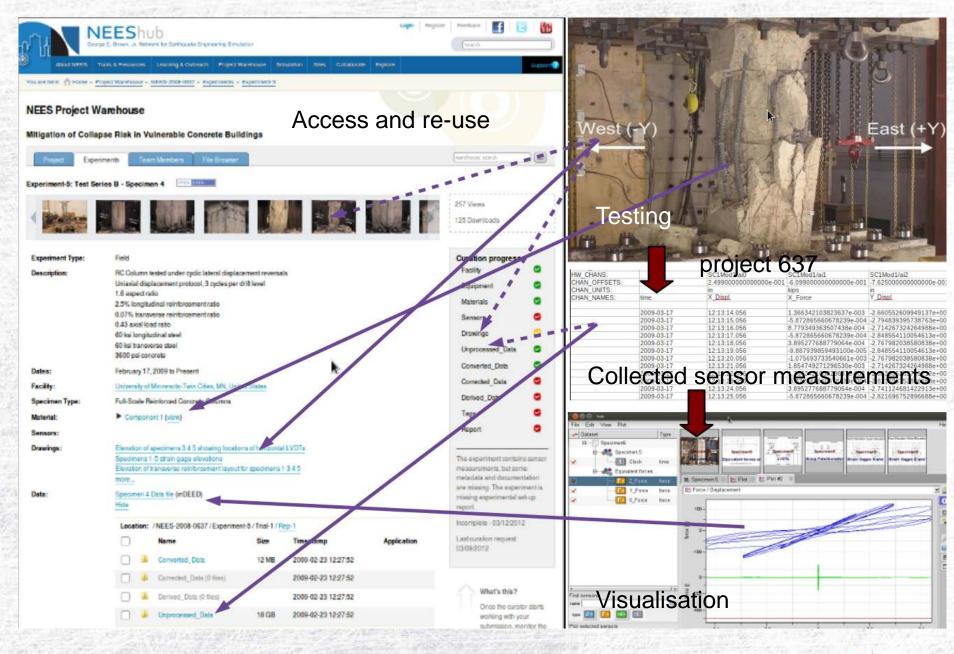
- Instrumentation plan
- Sensor metadata
- Data



Content - Metadata

- Names of researchers
- Affiliated organization
- Description
- Title
- Dates
- Testing facility
- Equipment
- Material properties
- Type of test
- Proper location
- Adequate file format
- Sensors









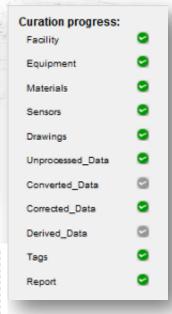
Curation - Path to SWAMP

- Straightforward (relatively)
- Way to
- Authorship
- Merit and
- Publication

Curation progress:

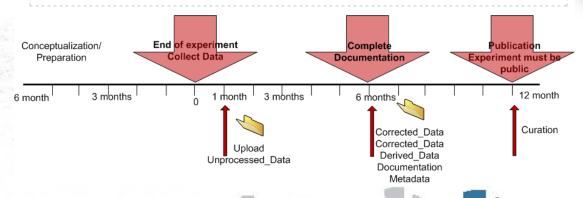
Experiments contain data and required metadata and documentation. The experimental setup reports and the final report were uploaded, as well.

Complete - 08/06/2012



Cite this work:

Clement Barthes, Catherine Whyte, Selim Gunay, Sangjoon Park, Donald Patterson, Shakhzod Takhirov, Bozidar Stojadinovic (2013). "Hybrid Simulation of the Seismic Response of Squat Reinforced Concrete Walls - Wall 2 Test", Network for Earthquake Engineering Simulation (distributor), Dataset, DOI:10.4231/D3125Q89N





Thank you.

Questions?

Comments?

Standa Pejša - spejsa@purdue.edu
NEEScomm Data Curator
Network for Earthquake Engineering Simulation