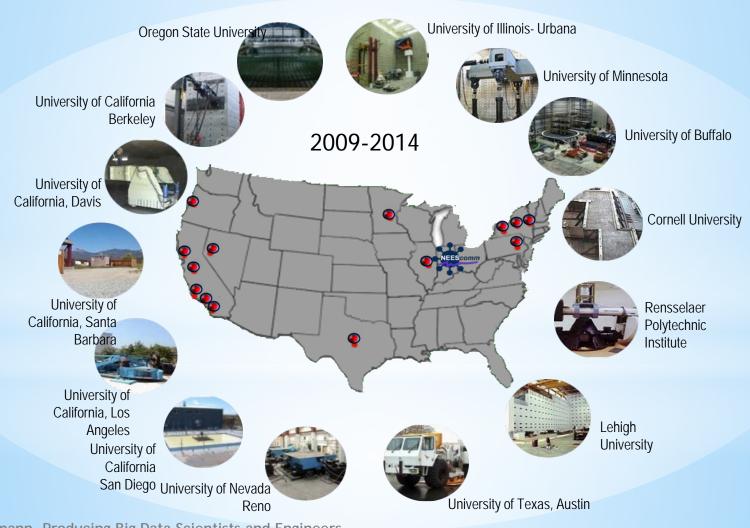
Producing the Next Generation of Data Scientists and Engineers

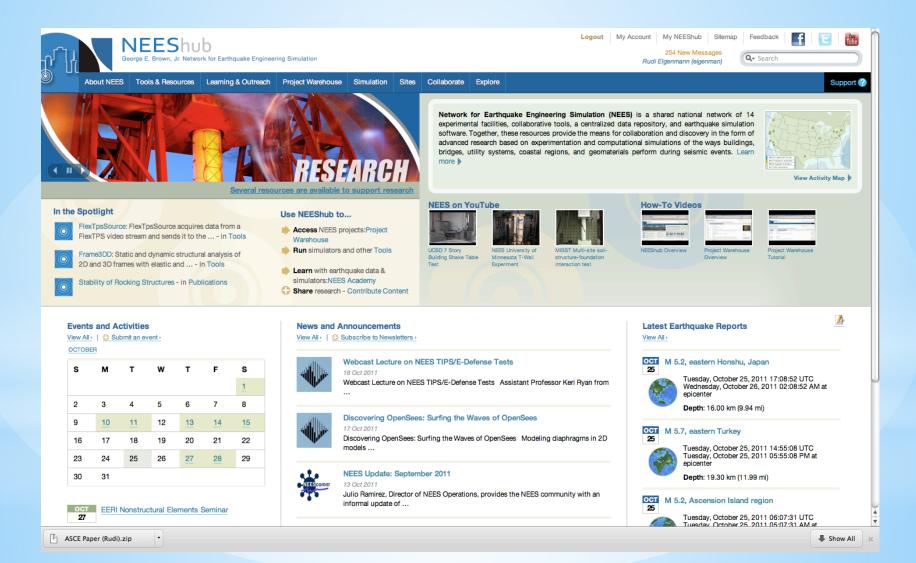
Rudi Eigenmann
School of Electrical and Computer Engineering
NEES Operations
Purdue University

The Context: NEEShub - nees.org

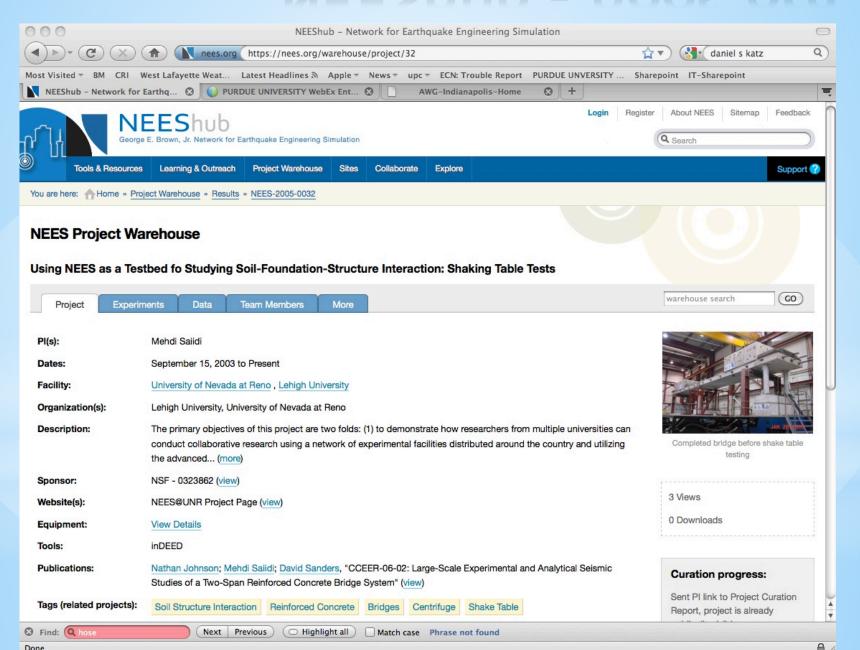
George E. Brown Jr. Network for Earthquake Engineering Simulation



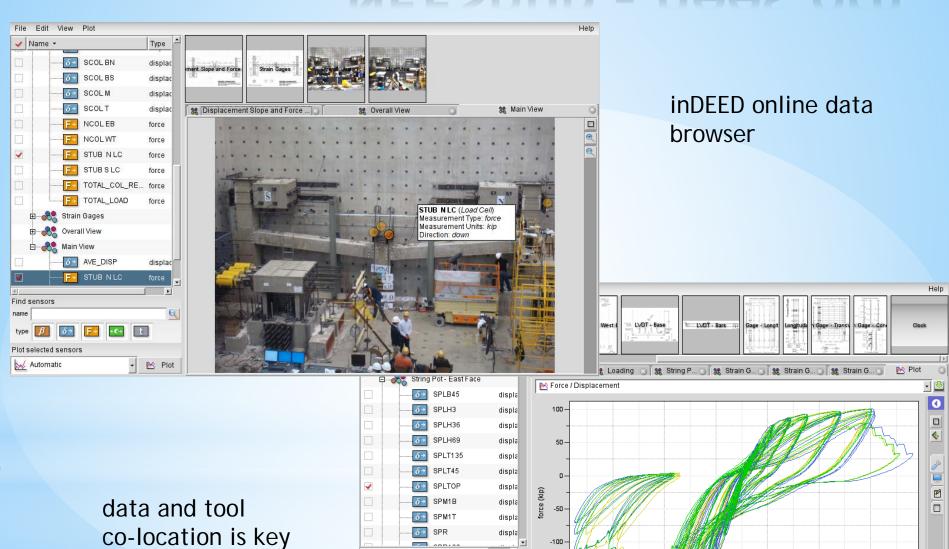
NEEShub - nees.org



NEEShub - nees.org



NEEShub - nees.org



M Plot

ϵ€→ t

-150

displacement (in)

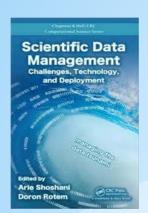
Find sensors

Plot selected sensors

Automatic

name

Rudi Eigenmann, Producing Big Data Scientists and Engineers

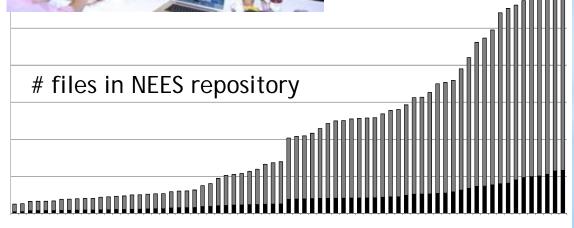


Data, Data, Data...





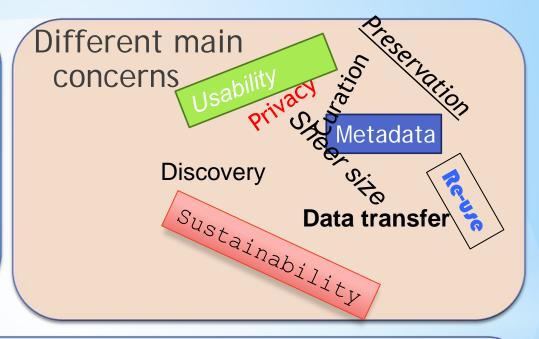




NEES Data by Comparison: Scope & Diversity of data

NEES has 14 sites, engineering data





Others: Data from earth & universe observation, biology, accelerators, medical, GPS, administrative info











How to Produce the Next Generation of Data Scientists and Engineers?

Purdue Interdisciplinary Program in

Big Data Science & Engineering

modeled after Purdue's Computational Science and Engineering program

Degree
Specialization in
CSE,CLS,BDSE
in addition to
PhD, MS degree

Interdisciplinary
Graduate Programs

CS&E
CLS

BDS&E

Core courses
Electives

*under development

BDS&E Courses

Enabling Technology

- * Computer Science, Computer Engineering: introduction to data management, machine learning and data mining, multimedia systems, pattern recognition and decision making processes, distributed computing systems, information retrieval
- * Statistics: statistical theory; statistical models; statistical methods; exploratory data analysis; statistical Machine Learning; computational algorithms for data analysis, interactive languages for data analysis
- * Libraries: Information organization; digital collection management; data ingest, archival, preservation, access; digital data curation; metadata & resource description

Data-based scientific methods

- * Geo Sciences: Advanced Spatial Ecology and GIS, Digital Remote Sensing and GIS, Geomatics;
- * Physics: Computational Physics, Quantum Computing, Particle Physics;
- * Civil Engineering
- * Biology
- * Agriculture
- * ...

New course developments needed!

New BDS&E Faculty Needed

Purdue Faculty Hiring Cluster in "Big Data"

enabling technologies:

Computer Science, Computer Engineering, Libraries

application domains:

Civil Engineering, Physics

Concluding Remarks

- *Big Data enables a fundamental new scientific method

 "the 4th pillar of science"
- *Producing the next generation of Big Data Scientists and Engineering is a challenge in itself
- *Purdue is creating a new interdisciplinary graduate program in Big Data Science&Engineering
- *A new generation of Big Data faculty is also needed

 Purdue is hiring in Big Data