ENHANCING EARTHQUAKE ENGINEERING RESEARCH THROUGH THE DESIGNSAFE CYBERINFRASTRUCTURE

E. Rathje\textsuperscript{1}, S. Brandenberg\textsuperscript{2}, J. Padgett\textsuperscript{3}, L. Lowes\textsuperscript{4}, G. Mosqueda\textsuperscript{5}, P. Arduino\textsuperscript{6}, M. Eslami\textsuperscript{7}, and N. Marafi\textsuperscript{8}

ABSTRACT

The DesignSafe cyberinfrastructure (\url{www.designsafe-ci.org}) is part of the NSF-funded Natural Hazard Engineering Research Infrastructure (NHERI) and provides cloud-based tools to manage, analyze, and understand data for natural hazards and earthquake engineering research. The DesignSafe Data Depot provides private and public disk space to support research collaboration and data publishing through a web interface. The DesignSafe Discovery Workspace provides cloud-based tools for simulation, data analytics, and visualization; as well as access to high performance computing (HPC). Examples of the use of DesignSafe in earthquake engineering research include: (1) integration of computational simulation, data analysis, and data publishing, (2) interactive data reports that use Jupyter notebooks that allows researchers to interrogate data interactively within the portal, (3) the publication of reconnaissance data using the DesignSafe Reconnaissance Portal and HazMapper App, and (4) the use of enhanced data models to meet user needs. These examples demonstrate how DesignSafe is enhancing research in natural hazards and earthquake engineering.

\textsuperscript{1} Professor, Dept. of Civil, Architectural, and Environmental Engineering, University of Texas, Austin, TX 78712 (email: e.rathje@mail.utexas.edu)
\textsuperscript{2} Professor, Department of Civil and Environmental Engineering, UCLA, Los Angeles, CA 90095
\textsuperscript{3} Associate Professor, Dept of Civil and Environmental Engineering, Rice University, Houston, TX 77005
\textsuperscript{4} Professor, Department of Civil and Environmental Engineering, University of Washington, Seattle, 98195
\textsuperscript{5} Professor, Dept. of Structural Engineering, Univ. of California, San Diego, CA 92093
\textsuperscript{6} Professor, Department of Civil and Environmental Engineering, University of Washington, Seattle, 98195
\textsuperscript{7} Graduate Student, Department of Civil and Environmental Engineering, UCLA, Los Angeles, CA 90095
\textsuperscript{8} Graduate Student, Department of Civil and Environmental Engineering, University of Washington, Seattle, 98195

Enhancing Earthquake Engineering Research through the DesignSafe Cyberinfrastructure

E. Rathje¹ S. Brandenberg² J. Padgett³ L. Lowes⁴ G. Mosqueda⁵ P. Arduino⁶ M. Eslami⁷ and N. Marafi⁸

ABSTRACT

The DesignSafe cyberinfrastructure (www.designsafe-ci.org) is part of the NSF-funded Natural Hazard Engineering Research Infrastructure (NHERI) and provides cloud-based tools to manage, analyze, and understand data for natural hazards and earthquake engineering research. The DesignSafe Data Depot provides private and public disk space to support research collaboration and data publishing through a web interface. The DesignSafe Discovery Workspace provides cloud-based tools for simulation, data analytics, and visualization; as well as access to high performance computing (HPC). Examples of the use of DesignSafe in earthquake engineering research include: (1) integration of computational simulation, data analysis, and data publishing, (2) interactive data reports that use Jupyter notebooks that allows researchers to interrogate data interactively within the portal, (3) the publication of reconnaissance data using the DesignSafe Reconnaissance Portal and HazMapper App, and (4) the use of enhanced data models to meet user needs. These examples demonstrate how DesignSafe is enhancing research in natural hazards and earthquake engineering.

Introduction

The DesignSafe cyberinfrastructure (www.designsafe-ci.org) has been developed as part of the Natural Hazards Engineering Research Infrastructure (NHERI) to enable and facilitate transformative research in natural hazards engineering, which necessarily spans across multiple

¹ Professor, Dept. of Civil, Architectural, and Environmental Engineering, University of Texas, Austin, TX 78712 (email: e.rathje@mail.utexas.edu)
² Professor, Dept of Civil and Environmental Engineering, UCLA, Los Angeles, CA 90095
³ Associate Professor, Dept of Civil and Environmental Engineering, Rice University, Houston, TX 77005
⁴ Professor, Dept of Civil and Environmental Engineering, University of Washington, Seattle, 98195
⁵ Professor, Dept of Structural Engineering, Univ. of California, San Diego, CA 92093
⁶ Professor, Dept of Civil and Environmental Engineering, University of Washington, Seattle, 98195
⁷ Graduate Student, Dept of Civil and Environmental Engineering, UCLA, Los Angeles, CA 90095
⁸ Graduate Student, Dept of Civil and Environmental Engineering, University of Washington, Seattle, 98195

disciplines and can take advantage of advancements in computation, experimentation, and data analysis. DesignSafe allows researchers to more effectively share, find, analyze, and publish data; perform numerical simulations and utilize high performance computing (HPC); and integrate diverse datasets. DesignSafe has been developed as a flexible, extensible, community-driven cyberinfrastructure and it embraces a cloud strategy for the big data generated in natural hazards engineering. It provides a comprehensive cyberinfrastructure that supports the full research lifecycle, from planning to execution to analysis to publication and curation.

The main components of DesignSafe that facilitate research are: (1) the Data Depot, a flexible data repository with tools for data management, curation, and publication, (2) the Discovery Workspace that allows simulation, data analytics, and visualization to be performed in the cloud and linked with the Data Depot, and (3) the Reconnaissance Portal that provides easy access to data collected during post-event reconnaissance efforts. This paper provides a summary of these components and examples of how DesignSafe is being used in earthquake engineering research.

DesignSafe Components

The Data Depot is the central shared data repository that supports the full research lifecycle, from data creation to analysis to curation and publication. Researchers have access to the private “My Data” space, the semi-private and collaborative “My Projects” space, and the “Published” space for curated and publicly available data. The Data Depot provides an intuitive interface to facilitate interacting with the data. Upload/download of data is streamlined through a range of interactive and automated options for both single file and bulk transfer, including drag and drop file upload, federation with existing cloud data services (e.g. Box, Dropbox, Google Drive, Globus), and command line interfaces that can be automated by power users.

Data curation services are provided to all users in DesignSafe. Curation involves organizing data and gathering the documentation that is needed for its use now and in the future. DesignSafe provides the tools and resources required to fully curate complex datasets. These tools have been developed to handle the unique characteristics of different types of datasets, specifically Experimental data, Simulation data, Hybird Simulation data, Reconnaissance data, as well as Other data. DesignSafe has adopted a progressive approach to data curation, in which the research team can provide the curation information during the course of the research, and thus shares responsibility for the curation process. When initially uploaded, data may have limited or even no user-supplied metadata. As data progresses towards publication, the requirements for metadata increase and at publication the user may edit the metadata and complete the process of assigning Digital Object Identifiers (DOIs) and applying the appropriate license. On demand assistance from a curator is available to provide training and to guide users through their data curation and publication needs.

The Discovery Workspace is intended to be the preeminent place for researchers to analyze, visualize, and transform their data in the cloud, and perform simulations using the most sophisticated computational tools available. Within the web portal, the Discovery Workspace provides a wide variety of Apps that can access the files in the Data Depot. The Apps available within the Discovery Workspace is evolving over time, but the current deployment of tools
includes open source computational simulation tools (e.g., OpenSees, ADCIRC), as well as tools for both data analytics and visualization (e.g. MATLAB; Jupyter, jupyter.org; HazMapper). Jupyter is a particularly noteworthy component of DesignSafe. A Jupyter notebook is an electronic notebook that allows users to embed rich text elements, as well as computer code, graphs, and visualizations, within a single notebook that can be shared through the web. Over 40 different programming languages are supported in Jupyter, including Python and R, and MATLAB code can be easily converted, making Jupyter a versatile tool for research. Many of the tools within the Discovery Workspace have access to HPC resources, making it easy for researchers to employ these resources in their work. These HPC-enabled tools in the Discovery Workspace can be used without request of a specific DesignSafe HPC allocation. We can also provide command line access to HPC resources for more advanced researchers. Details regarding the DesignSafe HPC allocation policy can be found at https://www.designsafe-ci.org/rw/user-guides/allocations-policy/.

The Reconnaissance Portal is the main access point for data collected during the reconnaissance of windstorm and earthquake events. Reconnaissance activities produce diverse data, including infrastructure performance data (e.g., damage estimates, ground movements, coastal erosion, wind field estimates), remotely sensed data (e.g., photos, video, LIDAR point clouds, satellite imagery data), or human experiential data (e.g., social media data, societal impact data, survey or interview data). These diverse data types have different metadata requirements, but their use hinges on information regarding the location from which the data were collected. Therefore, the Reconnaissance Portal utilizes a mapping framework to display the natural hazard events for which reconnaissance data are available. The reconnaissance data is physically located in the Data Depot and accessible by analytics and visualization tools in the Discovery Workspace, but the Reconnaissance Portal provide improved discoverability of the data.

As of March 2018, the DesignSafe cyberinfrastructure has 1,886 registered users, 40 TB of published data are available in the Data Depot (which includes the data previously published under the NEES program), and 241 Projects have been created in which researchers are sharing, organizing, and curating data from across the coastal, earthquake, and wind engineering domains. Over 33 TB of data are available within the private “My Data” space and the semi-private “My Projects” space, and much of these data will eventually be made publicly available within the “Published” section of the Data Depot.

The Use of DesignSafe in Earthquake Engineering Research

After less than two years of deployment, DesignSafe is already influencing the research being performed in earthquake engineering. Jupyter notebooks are being used broadly to interact with both simulation and experimental data, and the Reconnaissance portal is actively being used by the reconnaissance community. Below are specific examples of how DesignSafe is being used by the research community.

Integration of Computational Simulation, Data Analysis, and Data Publishing

OpenSees and Jupyter notebooks in DesignSafe have been adopted for studying seismic earth pressures mobilized against retaining walls. M.G. Durante (UCLA postdoctoral scholar) ran
thousands of continuum simulations using the OpenSees deployed in the Discovery Workspace to analyze vertically inhomogeneous elastic soil retained by rigid and flexible walls and to develop solutions for the Winkler stiffness intensity for these systems. The continuum solutions were pseudostatic, involving horizontal body forces. The Winkler stiffness intensity expressions were then implemented into a Jupyter notebook solution coded in Python for computing time histories of seismic earth pressure based on a user-input surface ground motion (Figure 1). A graphical user interface was developed using the ipywidgets Python library (Figure 1). The Jupyter implementation runs very quickly, and is significantly more efficient than running dynamic continuum analyses in OpenSees. This study shows how computational tools available in DesignSafe were used to complement each other to develop an efficient tool for computing seismic earth pressures.

The M9 Project at the University of Washington and the USGS (https://hazards.uw.edu/geology/m9/) is studying the impacts of a large-magnitude, megathrust earthquake on the Pacific Northwest (PNW). The severity of a magnitude 9 earthquake on the PNW is largely unknown, because no ground-motion recordings are available for such megathrust earthquakes in the region. DesignSafe has provided the M9 research team an HPC allocation on the Stampede supercomputer at the Texas Advanced Computer Center (TACC, https://www.tacc.utexas.edu/) and the team is using it to develop suites of simulated ground-motions for possible magnitude 9 rupture scenarios [2, 3]. Upon the completion of the project, the results of the ground-motion simulations will be published as a curated dataset in the DesignSafe Data Depot.

Figure 1. Jupyter notebook that uses Winkler stiffness intensity expressions to compute time histories of seismic earth pressure based on a user-input surface ground motion and general model properties
Saygili et al. [4] developed a Jupyter notebook that computes a seismic hazard curve for the sliding displacement of a slope given the yield acceleration ($k_y$) and natural period ($T_s$) of the slope, and ground motion hazard information in terms of the hazard curve for peak ground acceleration (PGA) and the associated magnitude and distance deaggregations. The Jupyter notebook is developed to interface directly with the U.S. Geological Survey Unified Hazard Tool such that the required ground motion hazard information can be input directly from the U.S. Geological Survey website.

**Interactive Jupyter Notebook Interfaces with Datasets**

Jupyter notebooks are starting to be adopted broadly to provide an interactive interface with datasets. Data reports are commonly static objects (e.g., PDF format) that include tables of sensor position and orientation, material properties, model geometry, testing sequence, and any other information necessary for users to make sense of the experiment(s). Static data reports suffer from many shortcomings, including but not limited to the following: (1) tables of information are often replicas of an official version of the information (often a spreadsheet), and researchers must remember to update the data report any time the official version changes, (2) hundreds of pages are often required to plot all of the sensor data, which is inconvenient for users who often seek a single sensor response, and (3) model sketches sometimes lack dimensions desired by users. Esmaelizadeh et al. [5] created an interactive digital data report using a Jupyter notebook that solves these shortcomings (Figure 2). Tables of information in the digital data report are rendered from the official spreadsheet at the time the notebook is executed, meaning that the data report is always consistent with the official version of the underlying data. A digital data plotter enables users to browse through the events and sensors, and quickly plot and interact with the data they are seeking (Figure 2). The model sketch was uploaded to an Autodesk 360 server, and subsequently embedded as an iframe object inside the Jupyter notebook, enabling users to interact with the model sketch and measure dimensions they desire. The notebook text and equations were written in Latex, enabling rich layouts. This Jupyter notebook allows easy and informative access to the data published in the Data Depot as part of an experimental project.

![Figure 2](image)

**Figure 2.** Interactive data report developed by [5] that includes rich text, equations, a digital data plotter, and embedded AutoCAD model sketches.

Eslami et al. [6] developed a notebook for visualization of test data from a large dataset of triaxial tests on sands published in the DesignSafe Data Depot. These tests were performed
specifically so that after post-processing of the data, yield surfaces could be mapped. The yield surfaces can be used to develop new constitutive models where the yield surface is measured from laboratory tests. The Jupyter notebook provides plots of stress paths and stress-strain curves of a sand tested at various confining pressures (Figure 3). Eslami [personal communication] is also developing a Jupyter notebook to provide access to a dataset containing about 60 undrained, Direct Simple Shear (DSS) tests on low-plasticity silt-clay mixtures at varying OCRs. The notebook is capable of plotting all the test data in time series, as well as stress-strain and stress paths, and also integrates the data from multiple tests to generate a curve of cyclic stress ratio (CSR) versus number of cycles to liquefaction (N). It is also capable of drawing critical state lines (based on results of multiple monotonic tests) at desired strain levels, and also generates undrained shear strength plots versus OCR and regresses normalized strength parameters.

![Figure 3](image)

**Figure 3.** Interactive tool to interface with data from 60 DSS tests on low-plasticity silt-clay mixtures.

**Use of DesignSafe in Reconnaissance Efforts**

Fall 2017 was an active time for natural hazards, with hurricanes Harvey, Irma, Maria and Nate, and the Mexico City earthquake all occurring within a 6-week window. These events provided an opportunity for the reconnaissance community to make use of various DesignSafe functionalities that facilitate activities during field deployments and data integration/publishing after field deployments. Datasets associated with each of the natural hazard events are available via the Reconnaissance Portal (Figure 4). Selection of a natural hazard event, either from the list on the left or the map on the right, takes the user to an event page that provides details of the event and links to available datasets.

Many of the field reconnaissance teams collected geo-tagged photographs and associated data of field observations of damage. The HazMapper App in the Discovery Workspace provides users a mapping interface into which they can import their damage photos, GPS tracks, and other geospatial data, annotate the map with markers and polygons, and save the map as a GeoJSON file.
(either locally or to the Data Depot). Reconnaissance teams made use of HazMapper to integrate their field observations and photos of damage (Figure 4), and share them with team members for ultimate publishing in the Data Depot and posting in the Reconnaissance Portal.

Figure 4. (a) DesignSafe Reconnaissance Portal and natural hazard events for which data are available, (b) Use of HazMapper App to integrate and share damage data/photos from hurricane Harvey.
During field reconnaissance teams made great use of the DesignSafe Slack tool that can be accessed through a web host (https://designsafe-ci.slack.com) or the Slack App. Slack is an online collaborative communication tool that represents a modern, highly capable and integrative user forum. Communication can take place publicly via organized, topical channels or privately through direct messages between individuals or small groups. Files can be shared easily through drag and drop, and all content is indexed for easy search. Reconnaissance teams posted to specific Slack channels associated with each reconnaissance event and posted annotated pictures in real-time from the field, which allowed researchers from around the country keep track of the activities and communicate with the field teams (Figure 5). Slack can also be used to interact with researchers about specific topics through channels dedicated to different topics, such as Jupyter (Figure 5).

Figure 5. Use of the DesignSafe Slack tool during reconnaissance field deployments.

Enhanced Data Models to Meet User Needs

The DesignSafe DataDepot provides flexible data models to support a wide range of data derived from experiments, simulation and reconnaissance activities, with some examples provided in the previous sections. NHERI Experimental Facilities including, but not limited to Lehigh University, University of California at San Diego and University of Florida, have or are expanding capabilities to conduct hybrid simulations for earthquake and wind loading. Hybrid simulations combine both experimental and computer simulation substructures and curating these data requires a data model with enhanced capabilities to capture the complexities and various configurations involved in hybrid simulations. In collaboration with the NHERI Experimental Facilities as well as the NHERI SimCenter, a hybrid simulation data model has been developed and is currently under implementation in the Data Depot. The primary components of the data model are shown in Figure 6. The complete hybrid model is first described as a Global Model including a description of the
substructuring approach and boundary conditions. There are two main approaches to the software architecture considered for a hybrid test and both can be captured with the proposed model. In the first approach, the main numerical model is considered as the Master Simulation Coordinator with substructures considered as elements within the finite element framework. It can include one or more experimental substructures. Simulation substructures are typically not included unless different software is used to can provide more advanced modeling capabilities for a set of particular elements. In the second approach, the hybrid model is considered an assembly of the substructures with the Master Simulation Coordinator having the primary role of solving the global system response at each time integration time step and does not include any simulation elements. The primary role of the coordinator in this case is to ensure equilibrium and compatibility are satisfied at the boundaries of the substructures. In both cases, the Master Simulation Coordinator also contains connectivity between the substructures. To describe the experimental and simulation substructures included in a hybrid simulation, the data model utilizes the same experimental and simulation data models already developed by DesignSafe. For example, the only difference between an experimental substructure in a hybrid simulation and a regular shake table test is the input source.

![Figure 6. Hybrid simulation data model.](image)

**Conclusions**

The future of natural hazards engineering research requires integration of diverse data sets from a variety of sources, including experiments, computational simulation, field reconnaissance. The DesignSafe cyberinfrastructure provides the functionalities that will enable transformative research in natural hazards engineering. By adopting a cloud strategy, DesignSafe allows for a fundamental change in the way that research is performed. It provides a comprehensive cyberinfrastructure that supports cloud-based data analysis and visualization, interactive access to data, as well as the full lifecycle of experimental, field, and computational research required by engineers to effectively address the threats posed to civil infrastructure by natural hazards. In particular, Jupyter notebooks are being used within DesignSafe to provide improved access and integration of experimental and simulation, the Reconnaissance Portal and HazMapper App are being used to improve reconnaissance activities and data sharing, and the DesignSafe Slack tool is facilitating a virtual community of researchers who can easily interface to improve their research. The DesignSafe cyberinfrastructure is available to the global natural hazard research community and account registration is free. We encourage researchers to join and explore the ways in which DesignSafe can be used in their research.
Acknowledgments

The DesignSafe project is financially supported by the National Science Foundation under grant CMMI-1520817. The development of DesignSafe also leverages grant ACI-1134872 for high performance computing, and grants ACI-1127210 and ACI-1450459 for the development of the Agave API. This support is gratefully acknowledged.

References


