**DesignSafe – Cyberinfrastructure for NSF Natural Hazards Engineering Research Infrastructure**

https://www.designsafe-ci.org/

Natural hazards engineering plays an important role in minimizing the effects of natural hazards on society through the design of resilient and sustainable infrastructure. The DesignSafe cyberinfrastructure has been developed to enable and facilitate transformative research in natural hazards engineering, which necessarily spans across multiple disciplines and can take advantage of advancements in computation, experimentation, and data analysis. DesignSafe allows researchers to more effectively share and find data using cloud services, perform numerical simulations using high performance computing, and integrate diverse datasets such that researchers can make discoveries that were previously unattainable. This white paper describes the design principles used in the cyberinfrastructure development process, introduces the main components of the DesignSafe cyberinfrastructure, and illustrates the architecture of the DesignSafe cyberinfrastructure.

A cyberinfrastructure is a comprehensive environment for experimental, theoretical, and computational engineering and science, providing a place not only to steward data from its creation through archive, but also a workspace in which to understand, analyze, collaborate and publish that data. Our vision is for DesignSafe to be an integral part of research and discovery, providing researchers access to cloud-based tools that support their work to analyze, visualize, and integrate diverse data types. DesignSafe builds on the core strengths of the previously developed NEEShub cyberinfrastructure for the earthquake engineering community, which includes a central data repository containing years of experimental data. DesignSafe preserves and provides access to the existing content from NEEShub and adds additional capabilities to build a comprehensive CI for engineering discovery and innovation across natural hazards. DesignSafe has been developed along the following principles:

**Create a flexible CI that can grow and change**. DesignSafe is extensible, with the ability to adapt to new analysis methods, new data types, and new workflows over time. The CI is built using a modular approach that allows integration of new community or user supplied tools and allows the CI to grow and change as the disciplines grow and change.

**Provide support for the full data/research lifecycle**. DesignSafe is not solely a repository for sharing experimental data, but is a comprehensive environment for experimental, simulation, and field data, from data creation to archive, with full support for cloud-based data analysis, collaboration, and curation in between. Additionally, it is the role of a cyberinfrastructure to continue to link curated data, data products, and workflows during the post-publication phase to allow for research reproducibility and future comparison and revision. 

**Provide an enhanced user interface**. DesignSafe supplies a comprehensive range of user interfaces that provide a workspace for engineering discovery. Different interface views that serve audiences from beginning students to computational experts allow DesignSafe to move beyond being a “data portal” to become a true research environment.

**Embrace simulation**. Experimental data management is a critical need and vital function of the CI, but simulation also plays an essential role in modern engineering and must be supported. Through DesignSafe, existing simulation codes, as well as new codes developed by the community and SimCenter, are available to be invoked directly within the CI interface, with the resulting data products entered into the repository along with experimental and field data and accessible by the same analytics, visualization, and collaboration tools.

**Provide a venue for internet-scale collaborative science**. As both digital data captured from experiments and the resolution of simulations grow, the amount of data that must be stored, analyzed and manipulated by the modern engineer is rapidly scaling beyond the capabilities of desktop computers. DesignSafe embraces a cloud strategy for the big data generated in natural hazards engineering, with all data, simulation, and analysis taking place on the server-side resources of the CI, accessible and viewable from the desktop but without the limits of the desktop and costly, slow data transfers.

**Develop skills for the cyber-enabled workforce in natural hazards engineering**. Computational skills are increasingly critical to the modern engineer, yet a degree in computer science should not be a prerequisite for using the CI. Different interfaces lower the barriers to HPC by exposing the CI’s functionality to users of all skill levels, and best of breed technologies are used to deliver online learning throughout the CI to build computational skills in users as they encounter needs for deeper learning.

The DesignSafe infrastructure provides a comprehensive environment for experimental, theoretical, and computational engineering and science, providing a place not only to steward data from its creation through archive, but also the workspace in which to understand, analyze, collaborate and publish that data. The CI can be described in terms of the services it provides or in terms of the technical components that enable those services.

DesignSafe is architected to comprise the following services and components as shown in the figure:

* **DesignSafe** front end web portal
* The **Data Depot**, a multi-purpose data repository for experimental, simulation, and field data that uses a flexible data model applicable to diverse and large data sets and is accessible from other DesignSafe components. The Data Depot includes an intelligent search capability that allows dynamic creation of catalogs of the held data in an easily understandable way, and that can search ill-structured data with poor or incomplete metadata.
* A **Reconnaissance Integration Portal** that facilitates sharing of reconnaissance data within a geospatial framework.
* A web-based **Discovery Workspace** that represents a flexible, extensible environment for data access, analysis, and visualization.
* A **Learning Center** that provides training and online access to tutorials.
* A **Developer’s Portal** that provides a venue for power users to extend the Discovery Workspace or Reconnaissance Integration Portal, and to develop their own applications to take advantage of the DesignSafe infrastructure’s capabilities.
* A foundation of **storage and compute** systems at the Texas Advanced Computing Center (TACC), to provide both on-demand computing and access to scalable computing resources.
* A **middleware layer** to expose the capabilities of the CI to developers, and to enable construction of diverse web and mobile interfaces to data products and analysis capabilities
* A marketplace of **Community Defined Interfaces;** the extension capability of the CI allows other projects to leverage DesignSafe to build an interface of their own choosing.

The CI development was initiated in July 2015 upon receiving the NSF award, and was first deployed May 2016. As of June 2017 we have more than 1,100 registered users spanning dozens of institutions around the world.