

# January 6: Brandi Davis-Dusenbery, Scalable, Collaborative, Reproducible, and Extensible Analysis of TCGA Data in the Cloud



## SYNOPSIS:

The Seven Bridges Cancer Genomics Cloud pilot is one of three pilot projects funded by the National Cancer Institute. The overarching goal of the project is to explore how co-localizing large genomics datasets, like The Cancer Genome Atlas, with dynamic compute infrastructure to analyze them, can make learning from these data faster, and ultimately enable precision medicine.

In this seminar we'll highlight four guiding principles that have driven development of the Seven Bridges CGC:

**Making data *available* isn't enough to make it *usable*:** We've built a dynamic query engine that allows fast search of more than 140 clinical and biospecimen properties to enable finding interesting TCGA data faster and easier. Importantly, data are immediately available for analysis at scale using both pre-defined and custom workflows.

**The best science happens in teams:** A fine-grained permissions model allows transparent collaboration; in a secure and compliant manner.

**Reproducibility shouldn't be hard:** Each analysis, including all parameters, files, and software versions is fully logged and can be perfectly replicated days or months later.

**The impact of TCGA is amplified by new data and tools:** Researchers can readily bring their own data, and their own tools to analyze alongside TCGA data. Native implementation of the Common Workflow Language (CWL) specification enables portability of tools and workflows to and from other CWL-compliant systems.

The seminar will include a demo of the system and interested researchers can visit [www.cancer-genomics-cloud.org](http://www.cancer-genomics-cloud.org) to get involved.

## Session details...BIO:

Brandi Davis-Dusenbery is the Scientific Program Manager for the Seven Bridges CGC. She received her Ph.D. in Biochemistry from Tufts University and completed her postdoctoral studies at Harvard University. She is passionate about enabling the use of biomedical data to ultimately improve patient care.

## SUMMARY:


**Topic:** Scalable, Collaborative, Reproducible, and Extensible Analysis of TCGA Data in the Cloud

**Speaker:** Brandi Davis-Dusenbery, Ph.D.

**Date:** Wednesday, January 6, 2016

**Time:** 11 AM – 12 PM ET

You are invited to listen to Dr. Davis-Dusenbery's presentation in Room 2W908 in the NCI Shady Grove Building on Medical Center Drive or via WebEx.

**Presentation:** A screen cast of the presentation will be available for viewing after the event on the [NCI CBIIT Speaker Series YouTube Playlist](#). 

[View the presentation slides.](#)

*About the NCI CBIIT Speaker Series:*

The National Cancer Institute (NCI) Center for Biomedical Informatics and Information Technology (CBIIT) Speaker Series is a bi-weekly knowledge-sharing forum featuring both internal and external speakers on topics of interest to the biomedical informatics and research communities. For additional information, including past speaker series presentations, visit the [CBIIT Speaker Series page](#).

Individuals with disabilities who need reasonable accommodation to participate in this program should contact the Office of Space and Facilities Management (OSFM) at 240-276-5900 or the Federal TTY Relay number 1-800-877-8339.