

# Bringing Genomic Data to the Clinic 2022: *A Virtual Hackathon Hosted by Carnegie Mellon University Libraries and DNAnexus*

Apply Here: [BGDC 2022](#)

The “BGDC Consortium” is pleased to announce the next round of Bringing Genomic Data to the Clinic Hackathon!

**March 9-11, 2022**

**The hackathon will be held as a virtual event.**

## Topics

Working groups of eight to ten individuals will be formed into four to six teams. These teams will build pipelines to analyze large datasets within a cloud infrastructure.

- Building cancer subsetting dashboards
  - Team 1: Automated Literature searches for subsetting specific disease types
    - We are particularly encouraging medical librarians to attend this one
  - Team 2: Finding additional datasets for disease subsetting and disease efficacy
  - Team 3: Extending a shiny dashboard for clinical diagnostics and therapeutic exploration
  - Team 4: An information theoretical framework for estimating confidence in subtyping disease and recommending medication
- Genevar3 -- medically relevant structural variants in genic regions
- Visualization of RNAseq-based variant calling

## Organization

After a brief organizational session, teams will spend three days addressing a challenging set of scientific problems related to a group of datasets. Participants will analyze and combine datasets in order to work on these problems. Throughout the four days, we will come together to discuss progress on each of the topics, bioinformatics best practices, coding styles, etc.

## Datasets

Datasets will come from public repositories, with a focus on datasets featured in literature about colorectal and lung cancer, short read datasets in the [sequence read archive](#) that have been ported to cloud infrastructure, as well as derivative contigs of the above.

# Products

All pipelines and other scripts, software, and programs generated in this hackathon will be added to a public GitHub repository designed for that purpose ([github.com/collaborativebioinformatics](https://github.com/collaborativebioinformatics)).

Manuscripts describing the design and usage of the software tools constructed by each team may be submitted to an appropriate journal such as the [F1000Research hackathons channel](#), BMC Bioinformatics, GigaScience, Genome Research or PLoS Computational Biology and/or [BioHackrxiv](#).

# Application

Participants will be selected based on the experience and motivation they provide on the form.

If you confirm your acceptance, please make sure it is highly likely you can attend, as confirming and not attending prevents other data scientists from attending this event. Please include a monitored email address, in case there are follow-up questions.

Note: Participants will need to bring their own laptop to this program. A working knowledge of scripting (e.g., Shell, Python, R) is useful but not necessary to be successful in this event. Employment of higher level scripting or programming languages may also be useful.

Applicants must be willing to commit to all three days of the event.

This will be a virtual event. To read about all of the benefits of hackathons and how our virtual hackathons have worked in the past, please check out this blog post: <https://blog.dnanexus.com/2021-09-09-the-hybrid-hackathons-of-the-future-now-with-librarians/>.

Also, note that the hackathon may extend into the evening hours each day, particularly if you are in a timezone that is not US-Eastern Time. Please make any necessary arrangements to accommodate this possibility.

There will be no registration fee or cost associated with participating in this event.

For more information, or with any questions, please contact Ben Busby / Melanie Gainey