

Cold Spring Harbor Laboratory

Biological Data Science Hackathon!

Assisted by NCBI

From October 24th to 26th, Cold Spring Harbor Laboratory (CSHL), with assistance from NCBI, will host a Biomedical Data Science hackathon immediately before the [Biological Data Science Conference](#) at CSHL. The hackathon will primarily focus on writing functional software for advanced bioinformatics analysis of next generation sequencing data and metadata, but also may include analysis of other types of data, such as images or other molecular measurements. This event is for students, postdocs and investigators or others already engaged in the creation of pipelines for genomic analyses from next generation sequencing data, imaging data or metadata*. The event is open to anyone selected for the hackathon, and willing to travel to CSHL*. Working groups of 5-6 individuals will be formed into four to five teams. These teams will build novel pipelines, tools, and visualizations to analyze large datasets within a cloud infrastructure. The potential subjects for this hackathon include structural variant identification & analysis, genome assembly, single cell transcriptome & epigenome analysis, ultrafast genomic mapping, data encryption, deep learning technologies, and image analysis. Please see the [application](#) for specific team projects.

* Attendees of the CSHL Biological Data Science will be given preference, if space is limited. Also, there will be a nominal fee for attendees, partly to cover refreshments during the events.

** Some projects are available to other non-scientific developers, mathematicians or librarians

Organization

After a brief organizational session, teams will spend three days analyzing 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.

Datasets

Most of the datasets will come from the public repositories, primarily those housed at the NCBI. During the course, participants will have an opportunity to include other datasets and tools for analysis. Please note, if you use your own data during the course, we ask that you submit it to a public database within six months of the end of the event if possible.

Products

All pipelines and other scripts, software and programs generated in this course will be added to a public GitHub repository designed for that purpose (github.com/NCBI-Hackathons). A manuscript outlining 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](#). Continued development of the technology after the hackathon is also encouraged.

Application

To apply, complete this [form](#) (approximately 10 minutes to complete). Applications are due September 16th, 2016 by 4 pm ET. Participants will be selected from a pool of applicants based on the experience and motivation they provide on the form. Prior participants and applicants are especially encouraged to reapply. The first round of accepted applicants will be notified on September 19th by 5 pm ET, and have until September 22nd at 9 am to confirm their participation. If you confirm, please make sure it is highly likely you can attend, as confirming and not attending bars 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 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. *No financial support for travel, lodging or meals is available for this event is currently available, although attendees will be notified if that changes.* Also note that the course may extend into the evening hours on Monday and/or Tuesday. Please make any necessary arrangements to accommodate this possibility.

Please contact ben.busby@nih.gov with any questions.