Application for Participation in the post-Biological Data Science NCBI-style CSHL Hackathon November 10-12, 2018!

From 10-12 November 2018, the NCBI will help out with a Data Science hackathon at Cold Spring Harbor Laboratory, immediately following the Biological Data Science Meeting! The hackathon will focus on genomics as well as general Data Science analyses including text, image and sequence processing. This event is for researchers, including students and postdocs, who have already engaged in the use of large datasets or in the development of pipelines for analyses from high-throughput experiments. Some projects are available to other non-scientific developers, mathematicians, or librarians.

The event is open to anyone selected for the hackathon and willing to travel to CSHL. Please note that this follows directly on the heels of the Biological Data Science meeting.

Working groups of five to six individuals will be formed into five to eight teams. These teams will build pipelines and tools to analyze large datasets within a cloud infrastructure. Example subjects for such hackathons include:

- Prediction of infection-prone metagenomic states
- Endogenous Retroviral Expression in Cancer
- Expression of Structural Variants in Sepsis
- An Online Bioinformatics Pipeline Design Engine
- Disease clustering from literature based on limited training data (phenotypic information)
- Graphical User Interface for Gene Expression calculated on the fly from raw data

Please see the application form for more details and additional projects. The project list will continue to evolve and will be updated on the application form.

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.

Datasets

Datasets will come from public repositories or will be supplied by the project lead. During the hackathon, participants will have an opportunity to include other datasets and tools for analysis. Please note, if you use your own data during the hackathon, we ask that you submit it to a public database within six months of the end of the event.

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/NCBI-Hackathons). 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.

Application

To apply, complete this form (approximately 10 minutes to complete). Initial applications are due Friday August 17th, 2018 by 9 pm ET, and will applications that come in after the deadline will be accepted on a rolling basis. Participants will be selected based on the experience and motivation they provide on the form. Prior participants and applicants are especially encouraged to apply. For each round, participants will be asked to confirm their attendance shortly after acceptance. If you confirm, 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 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.* Also note that the hackathon will extend into the evening hours on

Saturday and Sunday. Please make any necessary arrangements to accommodate this possibility.

** Special for this event: Cold Spring Harbor will be providing room and board for participants at standard conference rates, and registration can be added to the biological data science meeting. Attendees may arrange alternative lodging, but are encouraged to contact carr@cshl.edu if they have questions about pricing or logistics. **

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

Venue: Cold Spring Harbor Laboratory

Additional Projects: If you have an additional project you would like to see added to the form, please submit it here!