

OMB # 0925-0740
Expiration Date: 05/2019

Public reporting burden for this collection of information is estimated to average 10 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: NIH, Project Clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-0740). Do not return the completed form to this address.

NCBI is excited to host our first Women-led Hackathon, a collaborative biodata science event organized by women on the NIH Campus in Bethesda, Maryland!

NIH has a strong interest in enhancing the diversity of the scientific workforce, and women in particular are underrepresented in data science. This women-led NCBI initiative strongly encourages researchers, especially women, at any stage of their data science journey to apply for this inaugural event. Past hackathon participants have ranged from students and postdocs with a working knowledge of scripting (e.g. Shell, Python, R) to those already engaged in the use of large datasets or in the development of informatics tools, code, or pipelines.

Hackathon Logistics

The hackathon runs from 9 am – 5 pm each day, with an optional social event on the evening of the second day. Working groups of five to six individuals, with various backgrounds and expertise, will be formed into five to eight teams with an experienced leader. These teams will build pipelines and tools to analyze large datasets within a cloud infrastructure. Throughout the three days, we will come together to discuss progress on each of the topics, bioinformatics best practices, coding styles, etc.

There will be no registration fee associated with attending this event.

Note: Participants will need to bring their own laptop to this program. No financial support for travel, lodging, or meals is available for this 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.

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, Genome Research, or PLoS Computational Biology.

Application Process

Applications are due Monday April 22th, 2019 by 5 pm ET. Participants will be selected based on the experience and motivation they provide on the form. Prior participants and applicants are especially encouraged to apply. The first round of accepted applicants will be notified on April 24th by 5 pm ET and have until April 26 at 5 pm ET to confirm their participation. If you confirm, you must be willing to commit to all three days of the event, as confirming and not attending prevents other data scientists from attending this event.

Legal

Entrants retain ownership of all intellectual property rights (including moral rights) in the code submitted to as well as developed in the hackathon. Employees of the U.S. Government attending as part of their official duties retain no copyright in their work and their work is in the public domain in the U.S.

The Government disclaims any rights in the code submitted or developed in the hackathon.

Participants agree to publish the code and any related data in GitHub.

* 1. What is your name (first,last)

* 2. What is your preferred email address

* 3. What is your education level

- Post baccalaureate fellow
- Bachelors degree
- Masters degree
- Graduate student
- PhD
- Post doctoral fellow

Other (please specify)

4. What institution are you affiliated with

5. What is your github username

* 6. Have you previously applied to participate in an NCBI hackathon

- Yes
- No

* 7. Have you previously participated in an NCBI hackathon

- Yes
- No

* 8. Have you previously participated in any other hackathons

- Yes
- No

* 9. Do you have a working knowledge of shell scripting

- Yes
- No

* 10. How comfortable are you with these programming languages

	Not at all	Beginner	Intermediate	Advanced
Python	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
R	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Java	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
C/C++	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
JavaScript	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Perl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Django	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Matlab	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ruby	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Other (please specify)

* 11. How comfortable are you with:

	Not at all	Beginner	Intermediate	Advanced
AWS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
GCP	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Docker	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Singularity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Git	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

* 12. Please tell us in 2-3 sentences what you work on

* 13. Please tell us in 2-3 sentences why you want to participate

* 14. Here are Some Approximate Project Titles (which are expected to evolve); please pick the top three you are interested in working on

- An open store for variant and gene prioritization tools
- Variable Tracking and Schema Capturing to make Biomedical Research Data 'FAIR'
- Molecular language: discovery of cell-to-cell communication molecules from RNAseq data
- Design of ICD-9 to 10 conversion function for the R package 'icd'
- dsVirus variant discovery and annotation pipeline
- Hiding in plain sight - unannotated structural variants in public genomic data sets
- New projects as they arise
- Other (please specify)

Done