Allie Hajian – Science Writer
Data Science Platform
Broad Institute of MIT/Harvard

A open-source, cloud-native platform for biomedical researchers

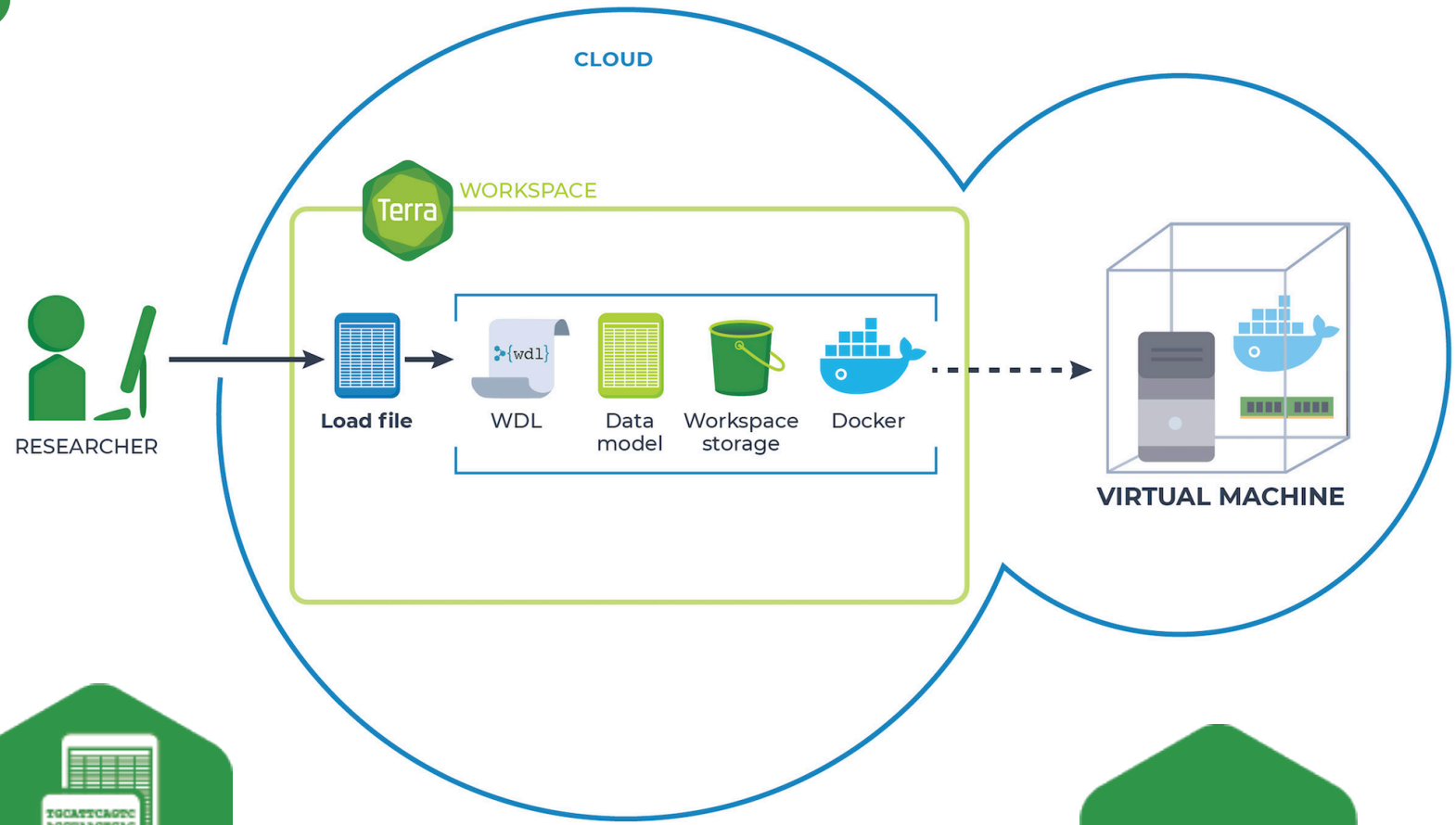
One place to

- Access data
- Run analysis tools
- Collaborate

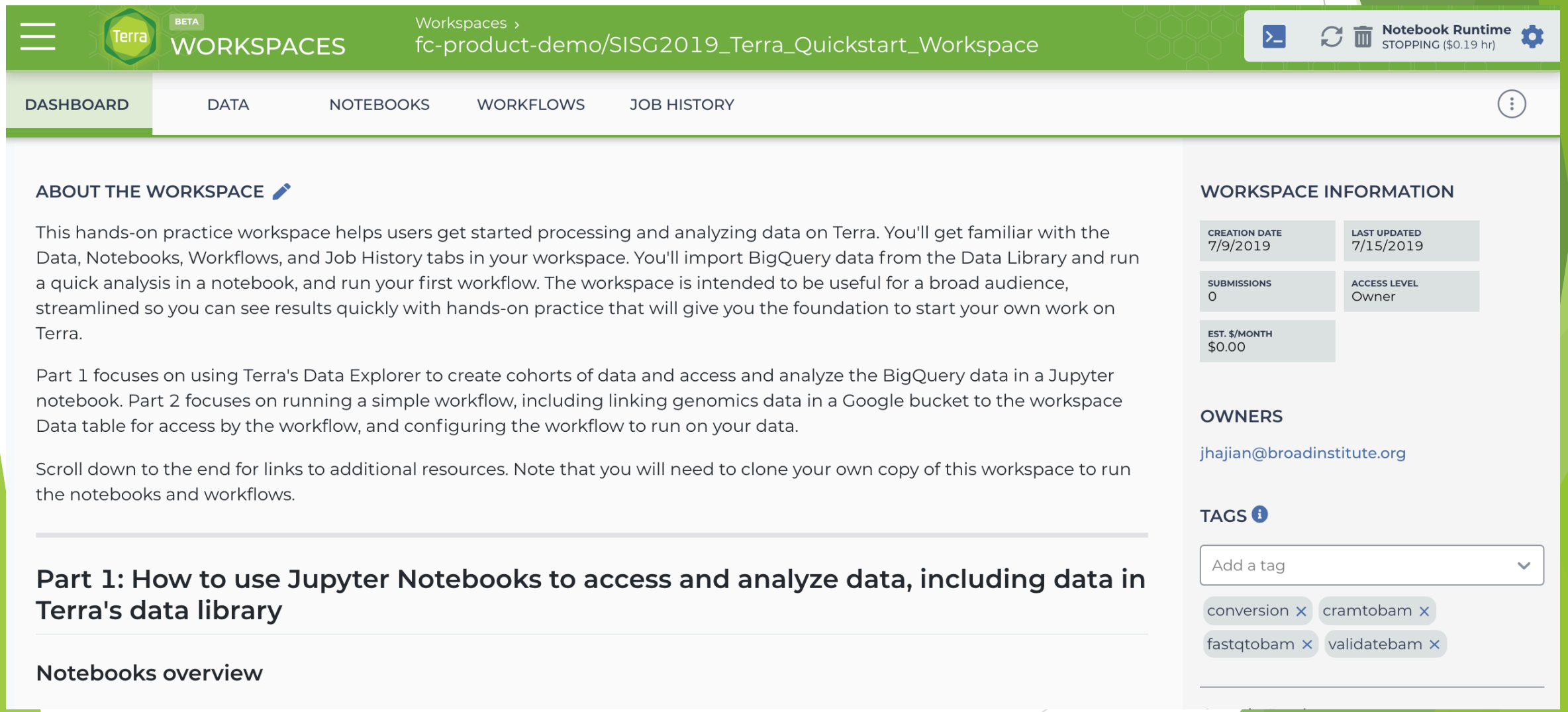
Designed with built-in security so you can focus on science

MODULAR	Comprised of functional components with well-specified interface
COMMUNITY FOCUSED	Created by many groups to foster a diversity of ideas
OPEN	Open-source licenses, software, arch to enable extensibility
STANDARDS BASED	Consistent with standards developed by coalitions such as GA4GH

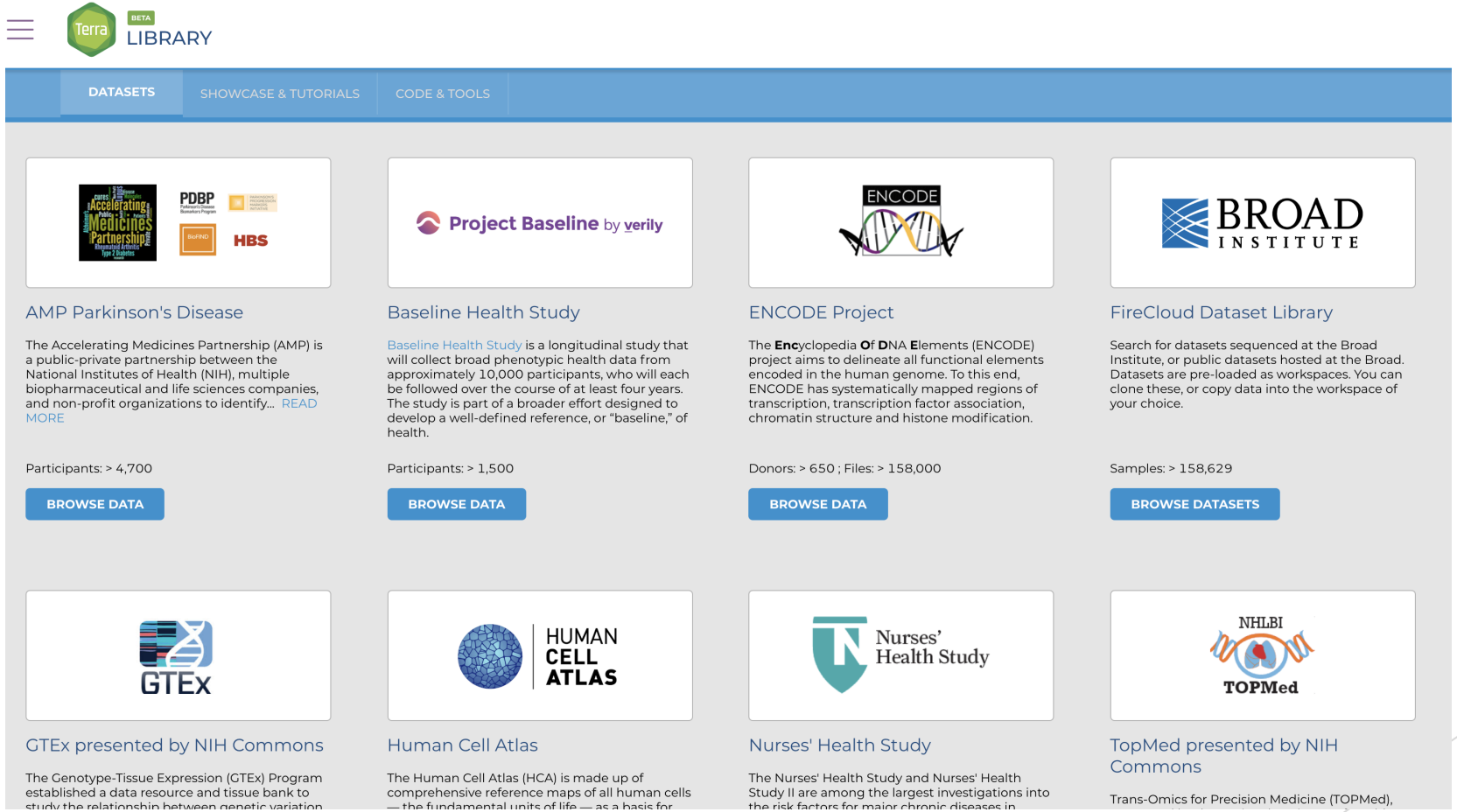
Core Functionality and organization



Workspaces for secure analysis and collaboration



Integrated Data Library, associated Google bucket



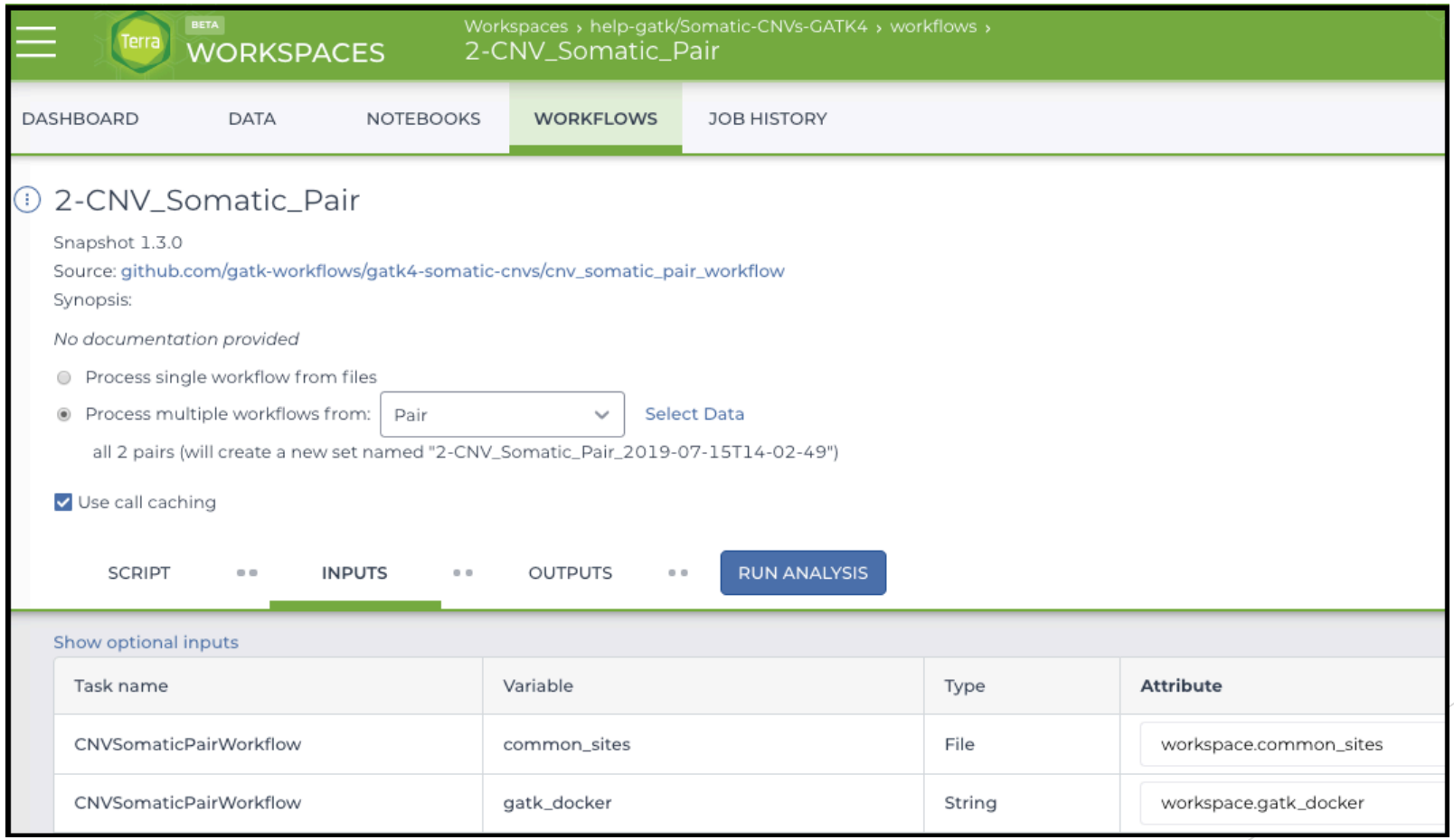
Analysis Tools including both interactive and batch processing

Interactive development environment for working with notebooks, code and data. Supports 100 programming languages with 1.7M public notebooks available

Scientific workflow engine designed for simplicity & scalability. Supports WDL and CWL, the two workflow languages adopted by GA4GH

Open-source, scalable framework for exploring and analyzing genomic data. Hail can generate variant/sample annotations, and perform variant, gene-burden, eQTL association analyses

Workflows to align, QC, call short variants per sample, joint-call across populations, and filter



Jupyter Notebooks for interactive visualization and analysis

"An open-source web application that allows users to create and share documents containing live code, equations, visuals, and narrative text"

Interactive Analysis

- Visualize and iterate in real time
- Couple with BigQuery to analyze data of any size or format in real time

Sharing

- Code and visualization tools, data, and narrative text in one package
- Lets your collaborators easily access and understand your results
- Open and interoperable standards

Our Ecosystem Partners

Terra is product of 

and

We are privileged to be part of the 
DATA
BIOSPHERE

Our goal is to build momentum for an *open, compatible, and secure* approach to data within the larger research community.

Acknowledgements

We are grateful to the  **team for their support.**

This walkthrough features a reproduction of the subset of analysis in:

Haas, ME et al. (2018) **Genetic Association of Albuminuria with Cardiometabolic Disease and Blood Pressure**. AJHG volume 103, issue 4, p461-473. *doi:10.1016/j.ajhg.2018.08.004*

This reproduction was developed in collaboration with members of the Kathiresan lab and specific thanks go to:

Mary Haas
Krishna Aragam
Sam Bryant
James Pirucello
Sek Kathiresan