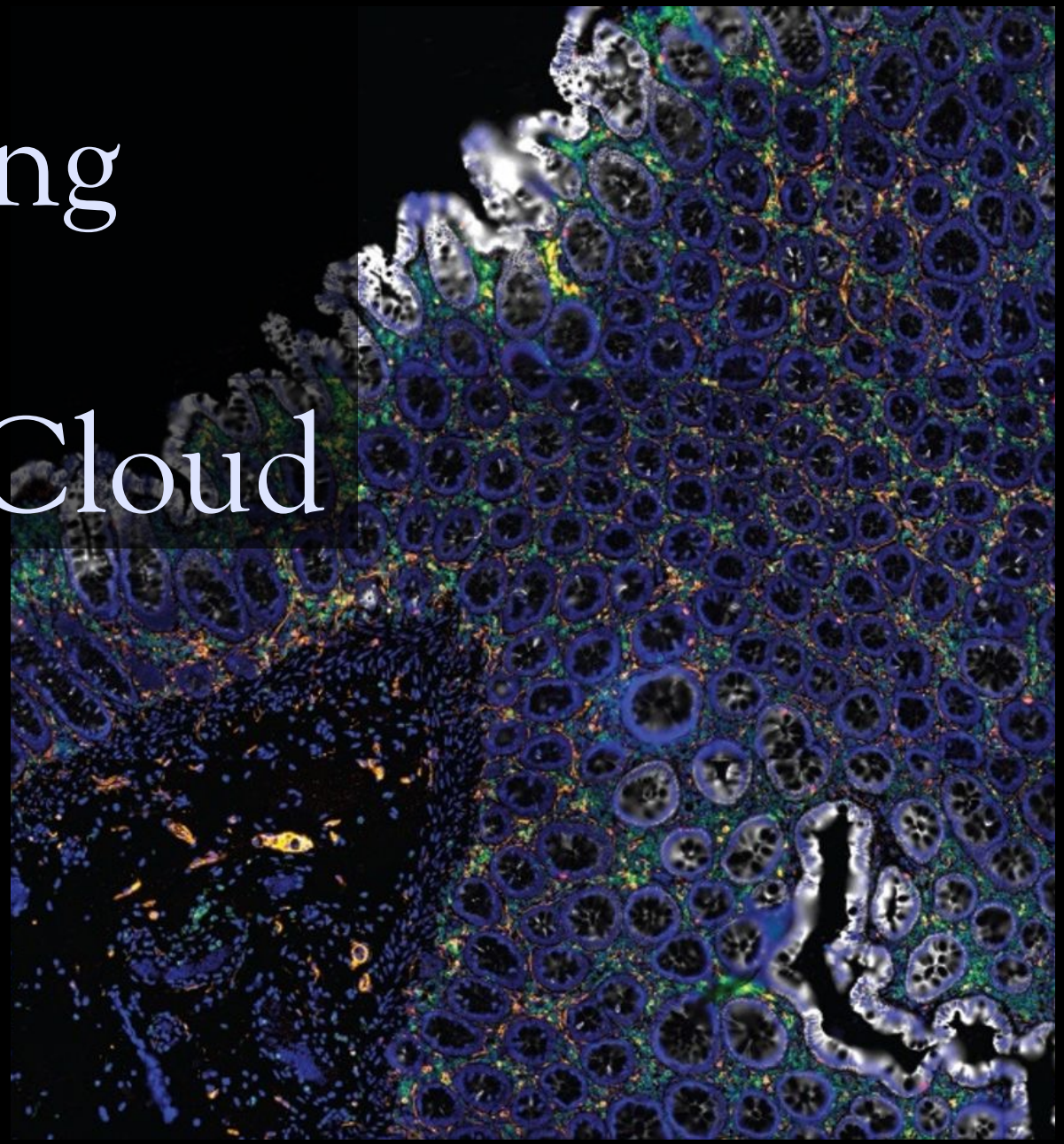
Bioinformatics Using The Seven Bridges Cancer Genomics Cloud



Seven Bridges Platform

Rowan Beck, PhD
March 27th, 2024

VELSERA

Agenda

Overview of the Seven Bridges Cancer Genomics Cloud, powered by Velsera
Features of the CGC
Live Demonstration

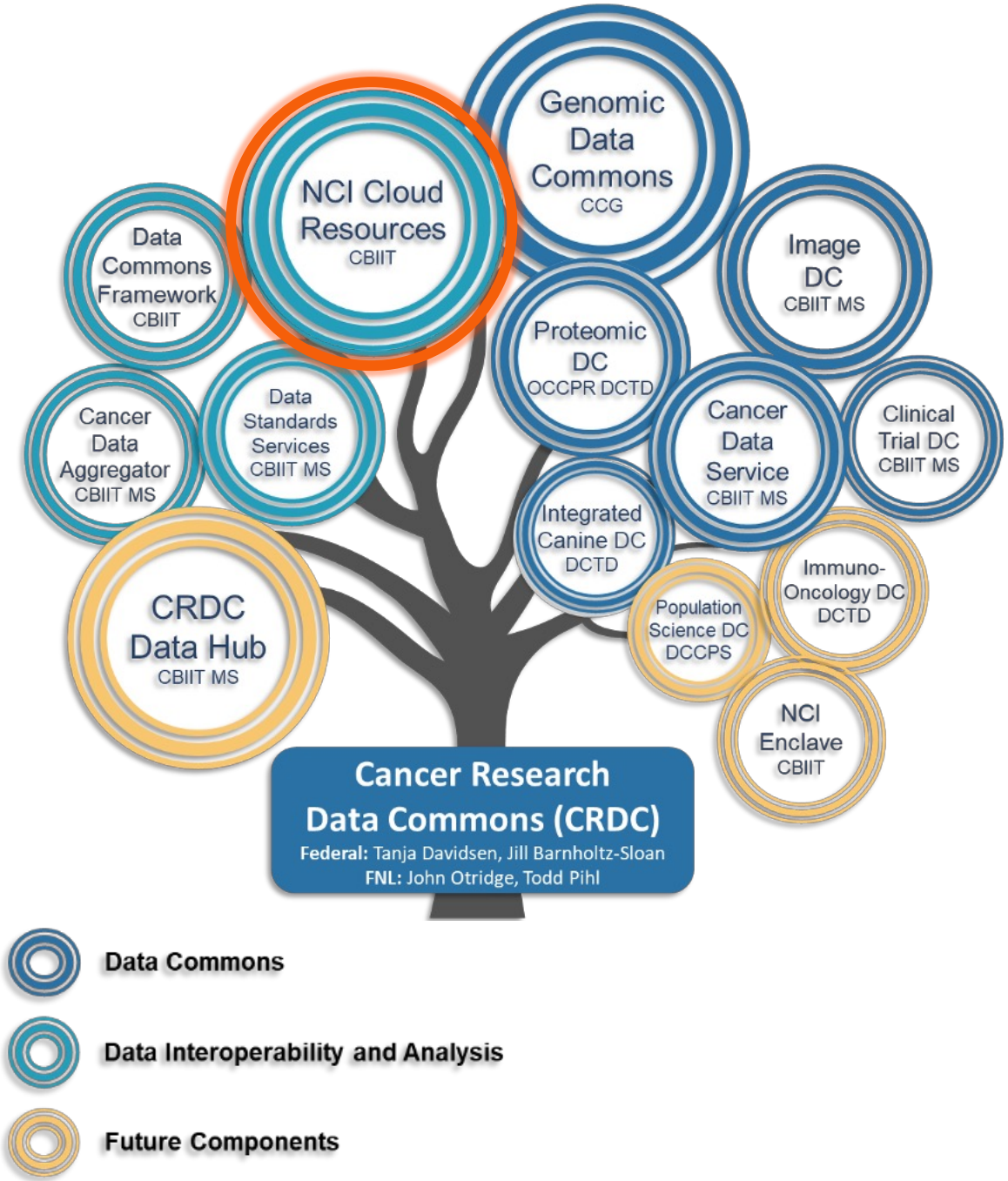
CANCER GENOMICS CLOUD
SEVEN BRIDGES

3+	1600+	900+	8000+	80000+
Petabytes Public Data	Years of Compute	Public Tools & Workflows	Users	User-Created Workflows

Provides powerful, yet easy to use interfaces to empower cancer researchers to draw new insights from petabyte scale data.

Stable, secure, and highly customizable cloud storage and computing platform

**Data Hub, CDA, & CTDC have not reached go-live*



VELSERA

Who are the CGC Users?

The CGC is designed to serve a wide range of scientists and users with varying skill sets

ADMINISTRATORS

- Manage and Control Users
- Monitor and Control Institutional Assets
- Manage and Monitor Projects
- Monitor and Control Costs
- Create Reports

BIOINFORMATICIANS

- Store, Manage, and Share Data
- Access Public and Proprietary Datasets
- Query, Build, and Investigate Cohorts of Interest
- Access Optimized Tools and Workflows
- Create, Optimize, Maintain, and Distribute New Tools and Workflows
- Create Push-button Automation Solutions
- Analyze Data at Scale with Tools and Workflows
- Conduct Interactive Exploratory Analyses
- Explore/Visualize Results and Gather Insights
- Easily Collaborate with Other Stakeholders
- Integrate with External Systems

BENCH SCIENTISTS

- Store, Manage, and Share Data
- Run Optimized Tools/ Workflows at Scale
- Conduct Defined Analyses via Push-button Solutions
- Investigate/Visualize Results
- Easily Collaborate with Other Stakeholders

CLINICIANS

- Conduct Validated Analyses via Push-button Solutions
- Query, Build, and Investigate Cohorts of Interest
- Create Reports
- Investigate/Visualize Results
- Easily Collaborate with Other Stakeholders

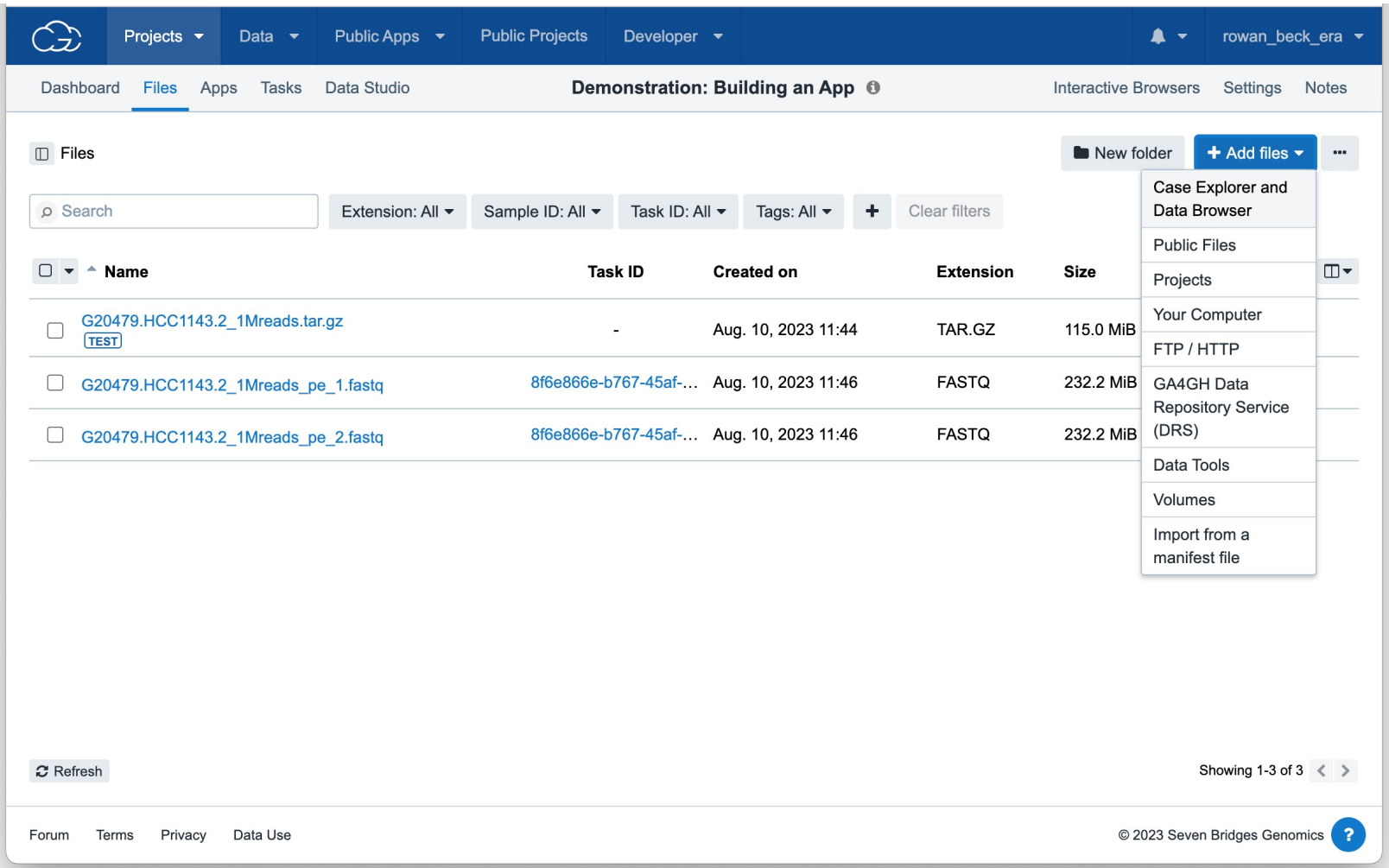
DEVELOPERS

- Create, Optimize, and Maintain New Tools and Workflows
- Create Push-button Automation Solutions
- Create Custom Interfaces for Specific Use Cases
- Distribute Proprietary Tools/ Workflows
- Integrate with Upstream/ Downstream Systems

Features of the CGC

VELSERA

Access Public Data, or Use Your Own



TCGA

ICGC

CPTAC

CANCER IMAGING ARCHIVE

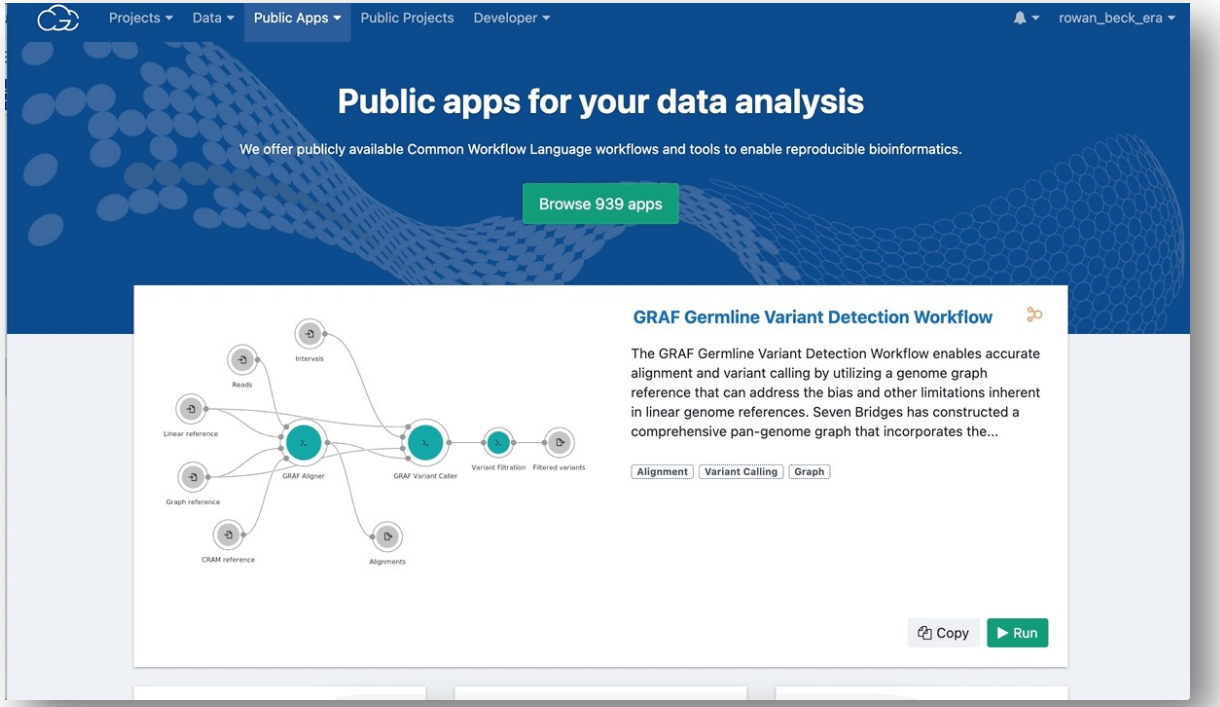
TARGET

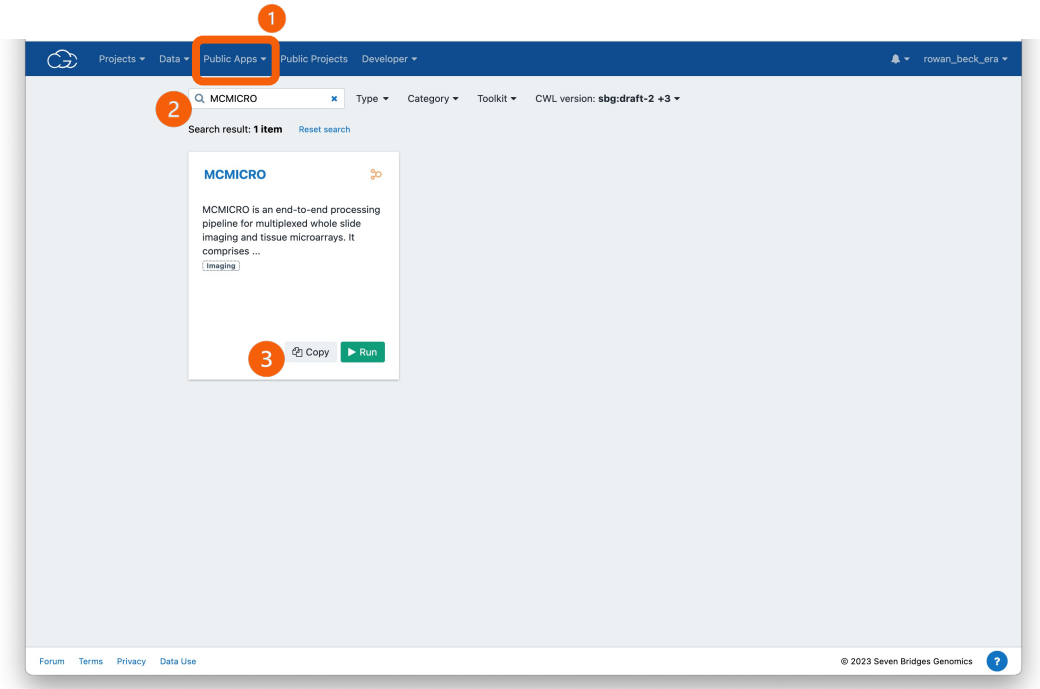
CCDI Childhood Cancer Data Initiative

Personal Genome Project

HTAN HUMAN TUMOR ATLAS NETWORK

Browse Hundreds of Tools and Workflows



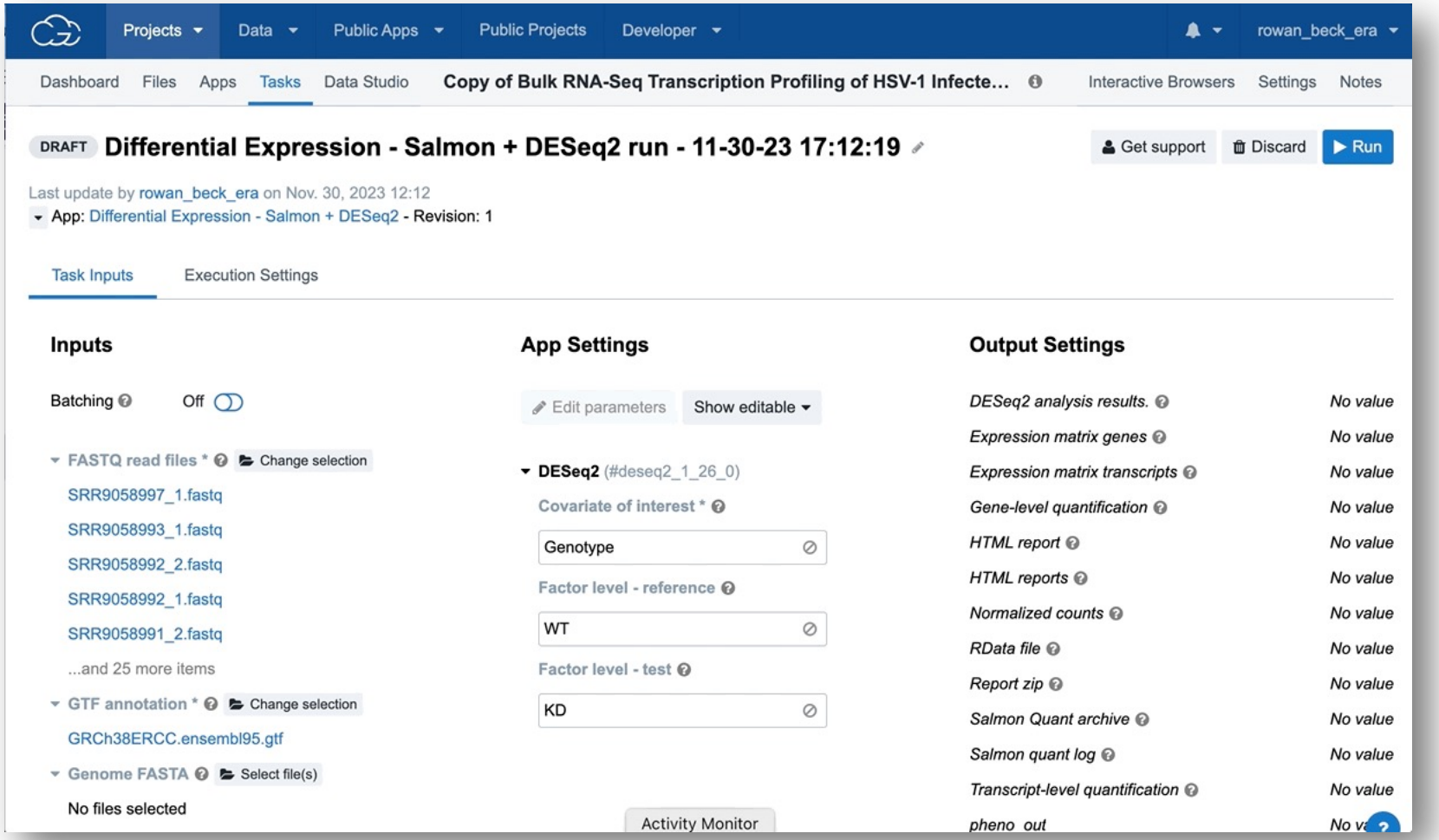


MCMICRO

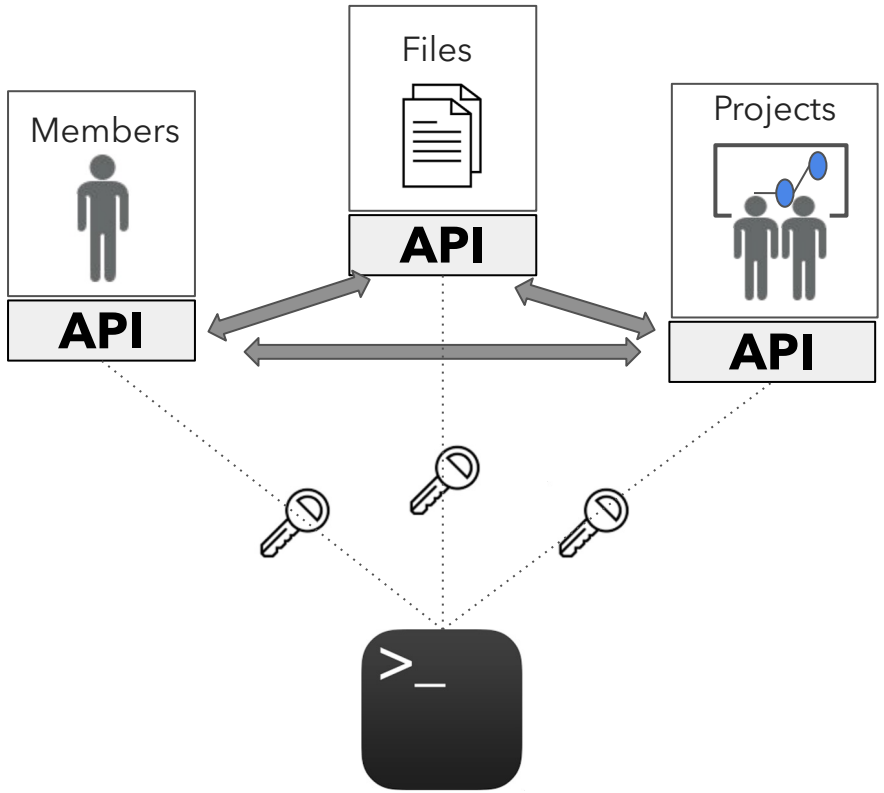
Enables the processing of multiplexed tissue images.

Transform whole-slide images into single-cell data using this simple workflow.

No Coding Required to Run an Analysis

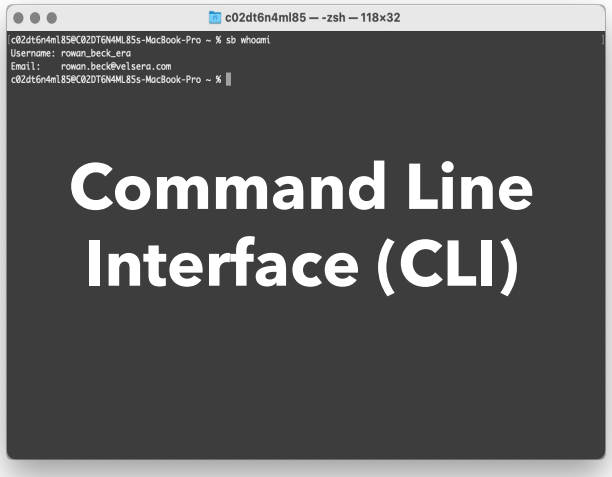


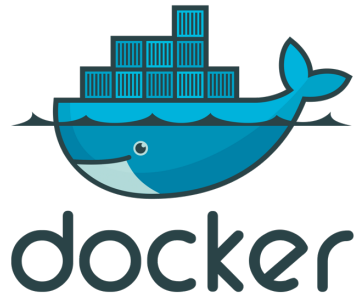
Coding Features



Example platform with Service Oriented Architecture and API access

- REST API
- API Bindings in Python, R, and Java





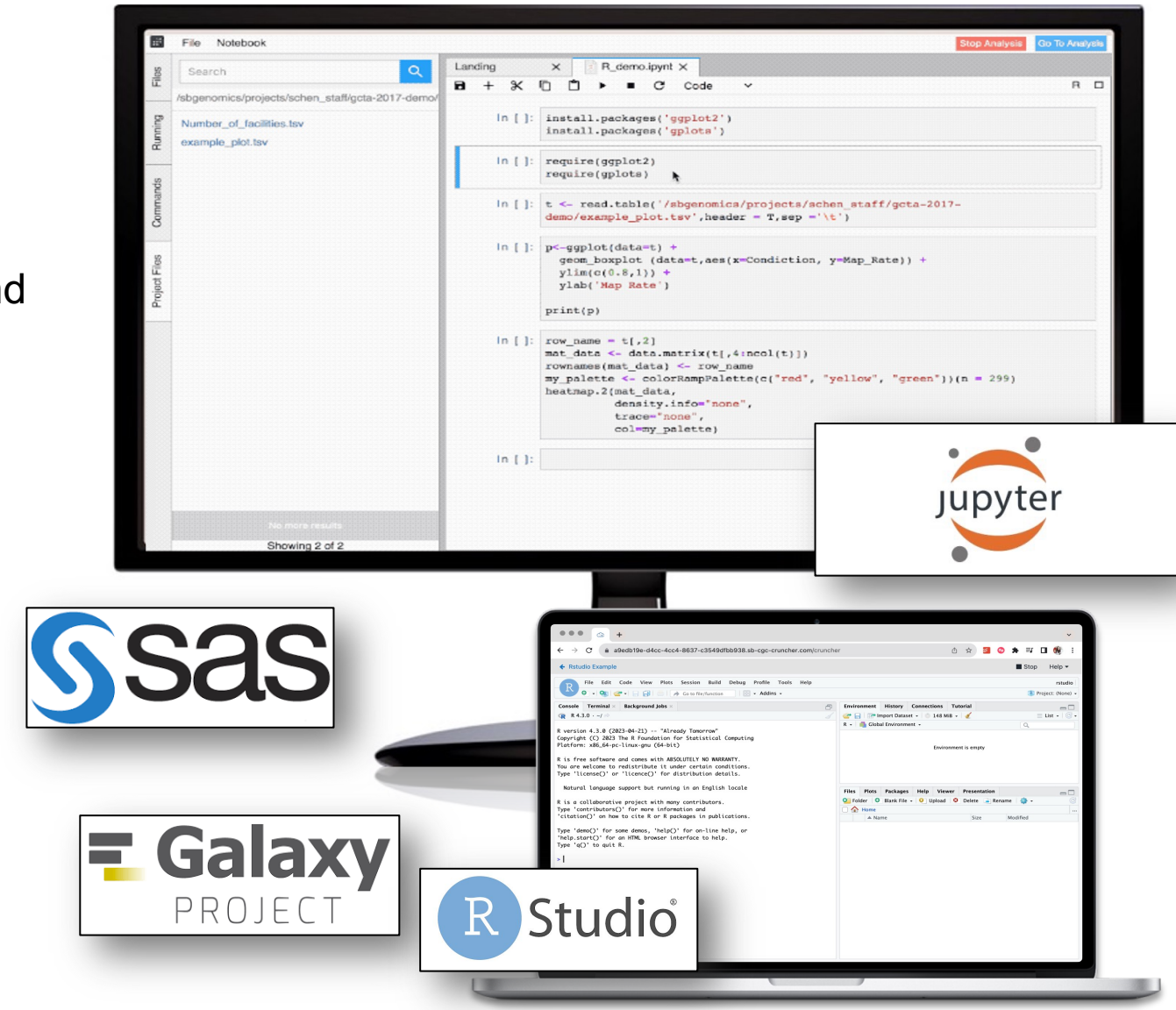
Integrated Custom Tertiary Analysis Tools

Data Science Workbench

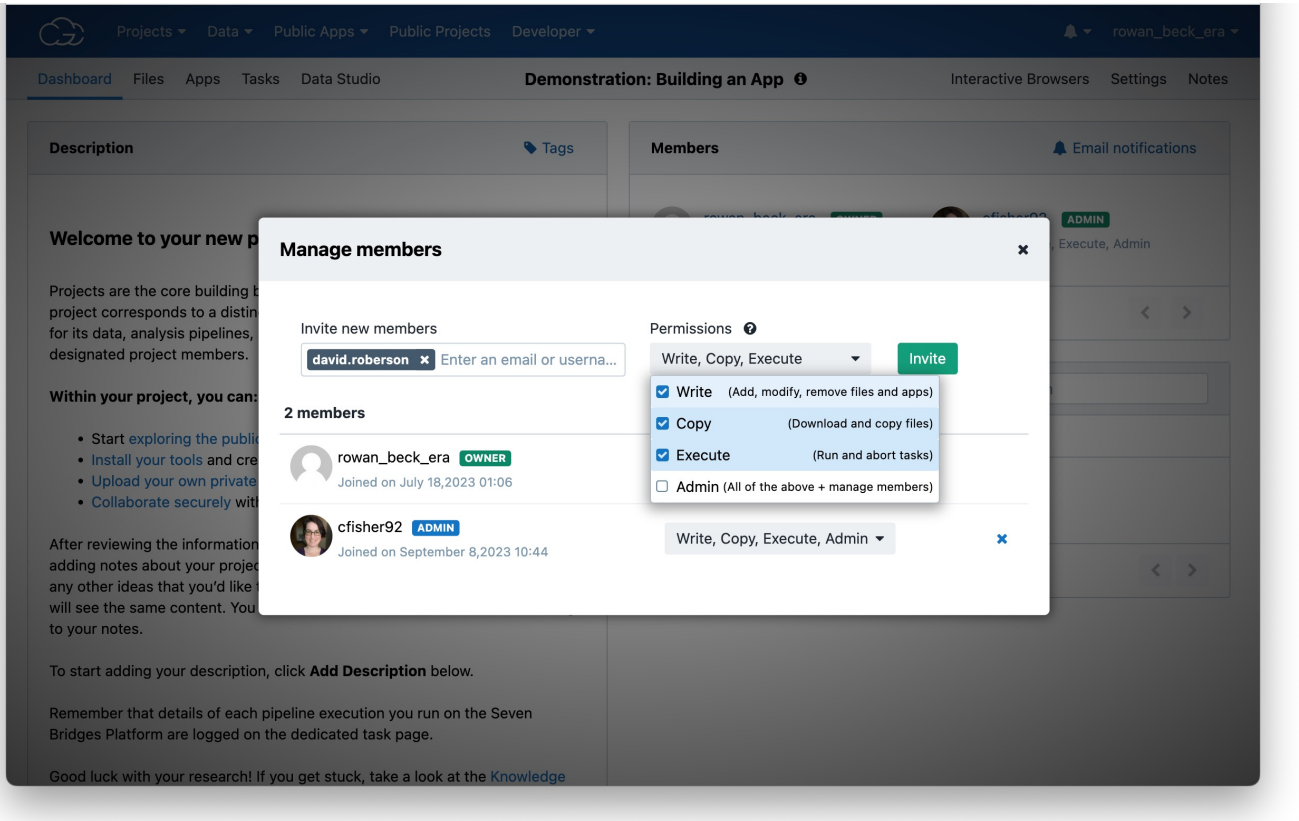
Derive new insights using interactive analysis environments with JupyterLab, RStudio, Galaxy, and SAS Studio environments. Create scripts and notebooks to record and share your analyses.

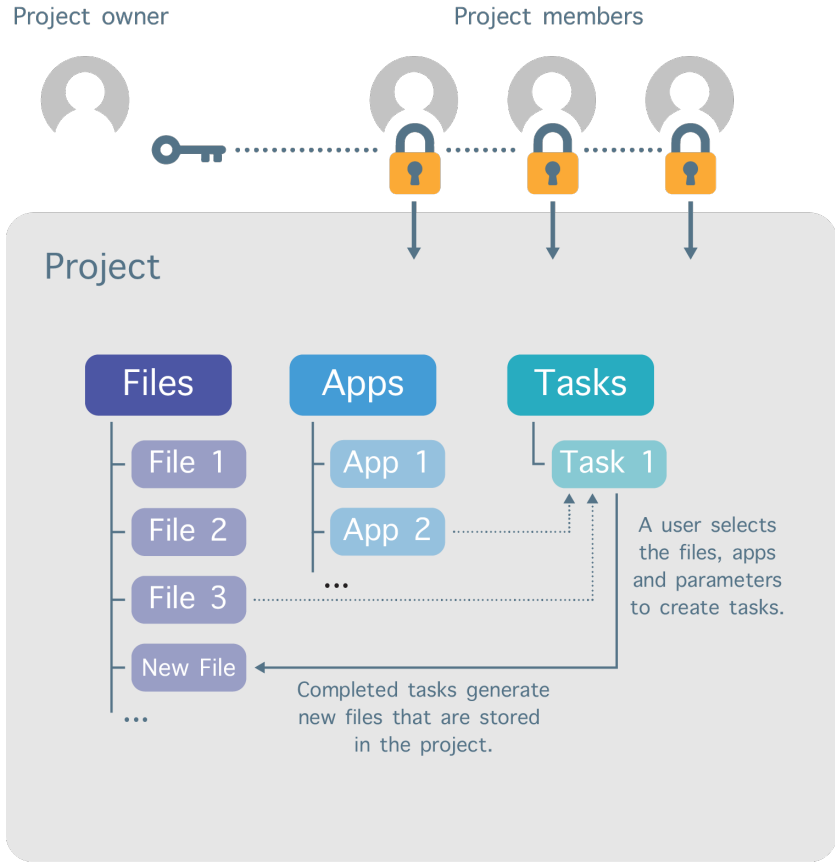
OHIF Image Viewer allows users to easily interact with DICOM images.

Open Health Imaging Foundation



Collaborating Has Never Been Easier





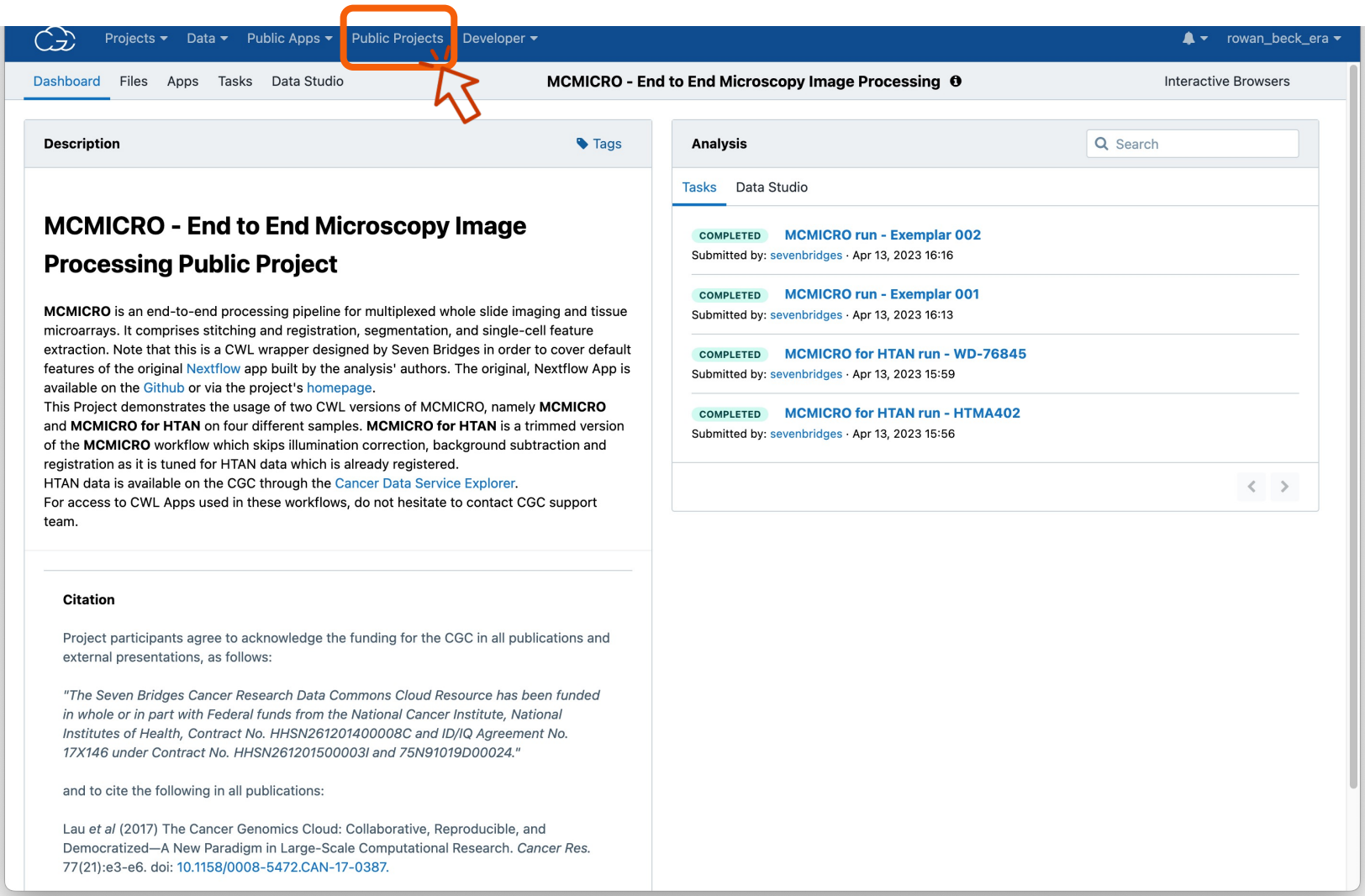
VELSERA

Collaborating Has Never Been Easier

Pre-loaded with

- input and output files
- pre-run tasks
- recommended settings

Detailed descriptions on what the workflow does and how to use the pipeline



VELSERA

Estimate Cloud Costs

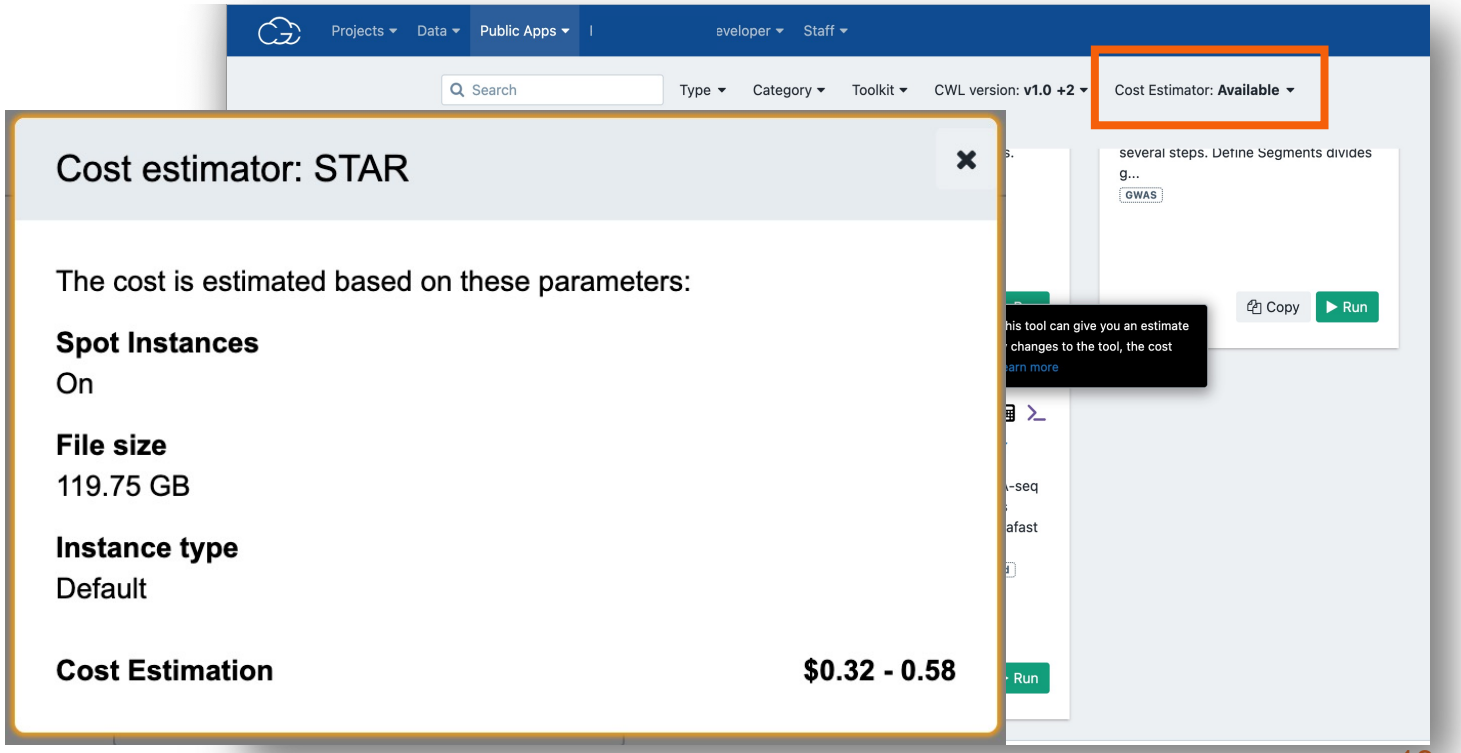
Performance Benchmarking

- Runtimes
- Task Costs
- Various file sizes

Experiment type	Input size	Paired-end	# of reads	Read length	Duration	Cost	Instance (AWS)
RNA-Seq	2 x 230 MB	Yes	1M	101	18min	$0.40	c4.8xlarge
RNA-Seq	2 x 4.5 GB	Yes	20M	101	30min	$0.60	c4.8xlarge
RNA-Seq	2 x 17.4 GB	Yes	76M	101	64min	$1.20	c4.8xlarge

Cloud Cost Estimator

- Available for a limited number of apps
- Create an estimate for the cost of your specific use case
- Compare workflows to save on computing costs



VELSERA

NCI Funding Is Available on the CGC

Diverse approaches and engagement strategies tailored to community needs

Pilot Credit Funds

- **$300 of cloud credits**
- Free for new CGC users
- Easy to request when signing up
- Fast approval

NIH NATIONAL CANCER INSTITUTE

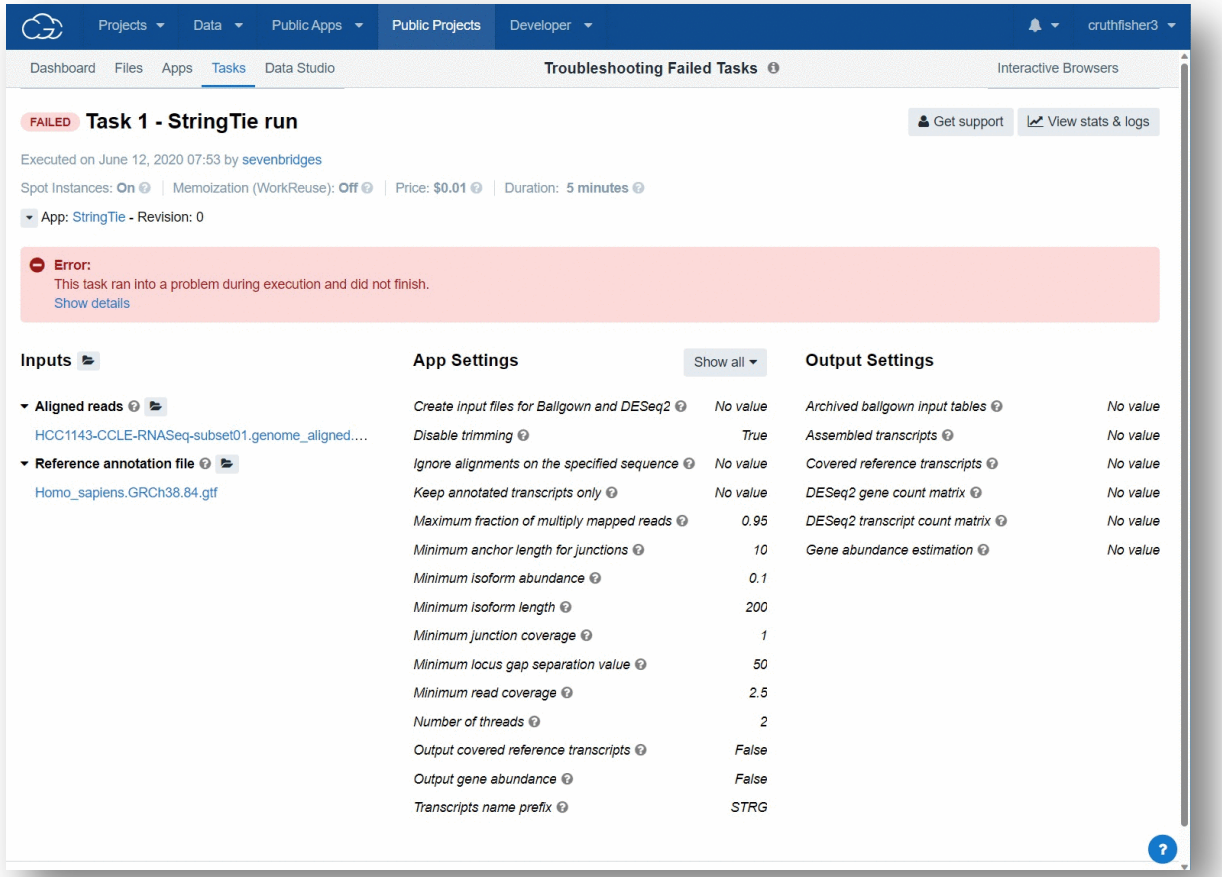
Collaborative Projects

- Cost estimation, optimization, and planning support
- Great for researchers new to bioinformatics and cloud approaches
- **Up to $10k compute/storage costs**
- Fast, rolling applications
- To date > 60 projects

Get The Support You Need

Every Week:

- 10:00 am ET Tuesday
- 2:00 pm ET Thursday



Create an account and access **free credits** at: CancerGenomicsCloud.org

Every Week:

- 10:00 am ET Tuesday
- 2:00 pm ET Thursday

Stay in touch.

Cera Fisher, PhD
Community Engagement Manager
Cera.Fisher@velsera.com

Rowan Beck, PhD
Community Engagement Manager
Rowan.Beck@velsera.com

Zelia Worman, PhD
Director of Researcher Engagement and Education
Zelia.Worman@velsera.com

Thank you.

Learn more at
CancerGenomicsCloud.org

VELSERA