Outreach and Training

29OCT2021

Jeff Leek and Tiffany Miller

AnVIL

Overview

- Mission
- Awareness
- Asynchronous Support
- Synchronous Support
- Tailored Support
- User Research
- Goal

2:30-3:45 Session 2: Breakout rooms

Infrastructure Moderators: Ms. Karen L. Davis (RTI International) and Dr. Carolyn M. Hutter (NHGRI)		Outreach and training Moderators: Dr. Siddharth Pratap (Meharry Medical College) and Mr. Christopher Wellington (NHGRI)	
2:30-2:35	Moderator introduction	2:30-2:35	Moderator introduction
2:35-2:50	AnVIL presentation: *Dr. Jeremy Goecks (OHSU) and Dr. Benedict Paten (UCSC)*	2:35-2:50	AnVIL presentation: *Dr. Jeffrey Leek (JHU) and Ms. Tiffany Miller (Broad)*
2:50-3:35	Discussion	2:50-3:35	Discussion
3:35-3:45	Prepare breakout report	3:35-3:45	Prepare breakout report

Breakout room: Outreach and training

Moderators: Dr. Siddharth Pratap and Mr. Christopher Wellington

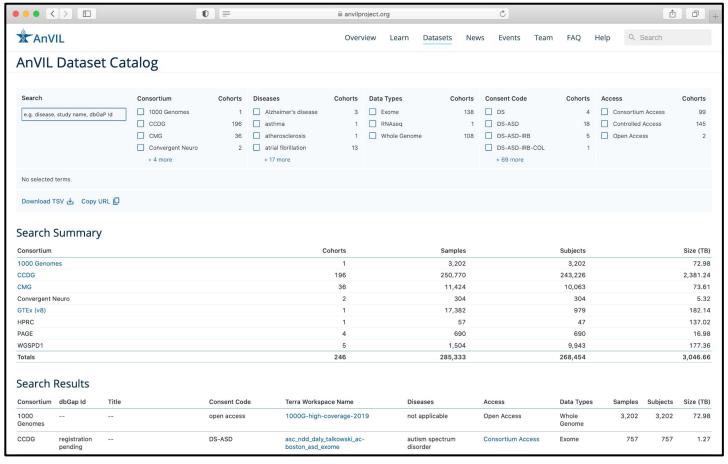
Dr. Cinnamon Bloss
Dr. C. Titus Brown
Dr. Carol Bult
Dr. John Kwagyan
Dr. Andrew Lee
Dr. Robert Meller
Dr. Peter Robinson
Dr. Sourav Roy
Dr. William (Bill) Southerland

Mission

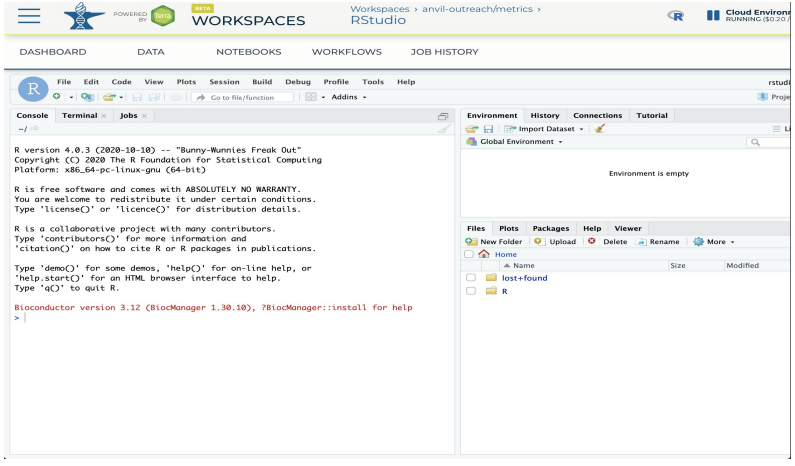
To develop scalable training, support, and incentives to make it easier for the scientific community to adopt and use the AnVIL Cloud Platform to share data, perform genomic and clinical analysis, & manage their computing.

Why AnVIL is awesome

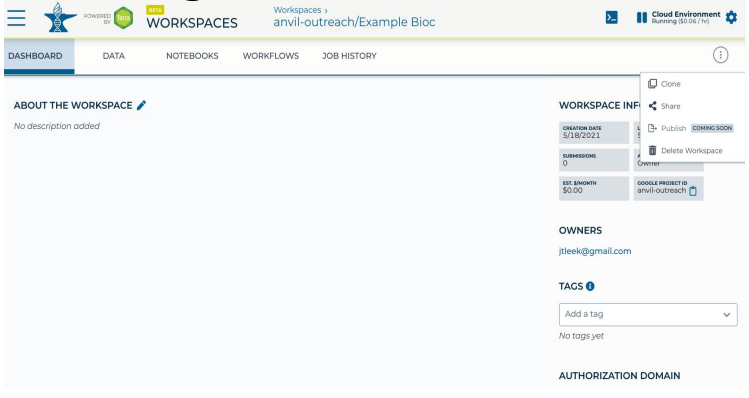
Data!



Computing!



Sharing!



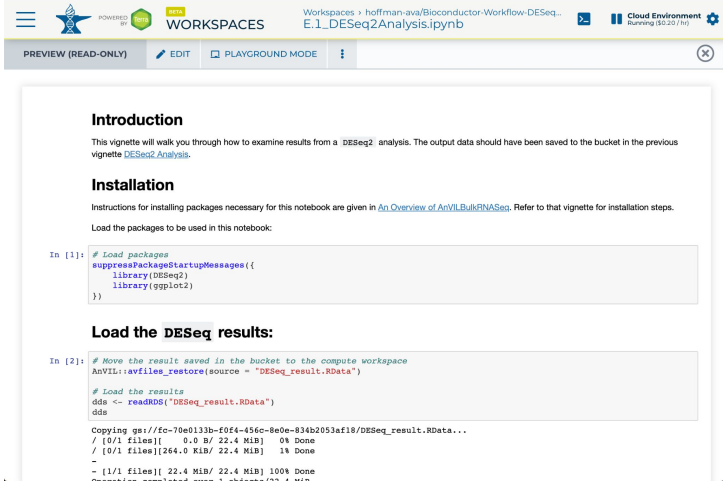
Scale!


YEEESSS!

Barriers to adoption and use of AnVIL

1. Making the community aware of AnVIL and its capabilities
    a. Getting AnVIL into users hands
    b. Keeping them updated about changes
2. Overcoming hurdles to cloud adoption
    a. Billing
    b. Security
    c. New way of thinking
3. Training people in using the cloud/AnVIL
    a. Using the cloud
    b. Managing people
    c. Managing data
    d. Managing computing
    e. Managing bills
4. Providing support to diverse communities
    a. Diverse in background/resources
    b. Diverse in applications
    c. Diverse in tools

AnVIL is “renting computers”



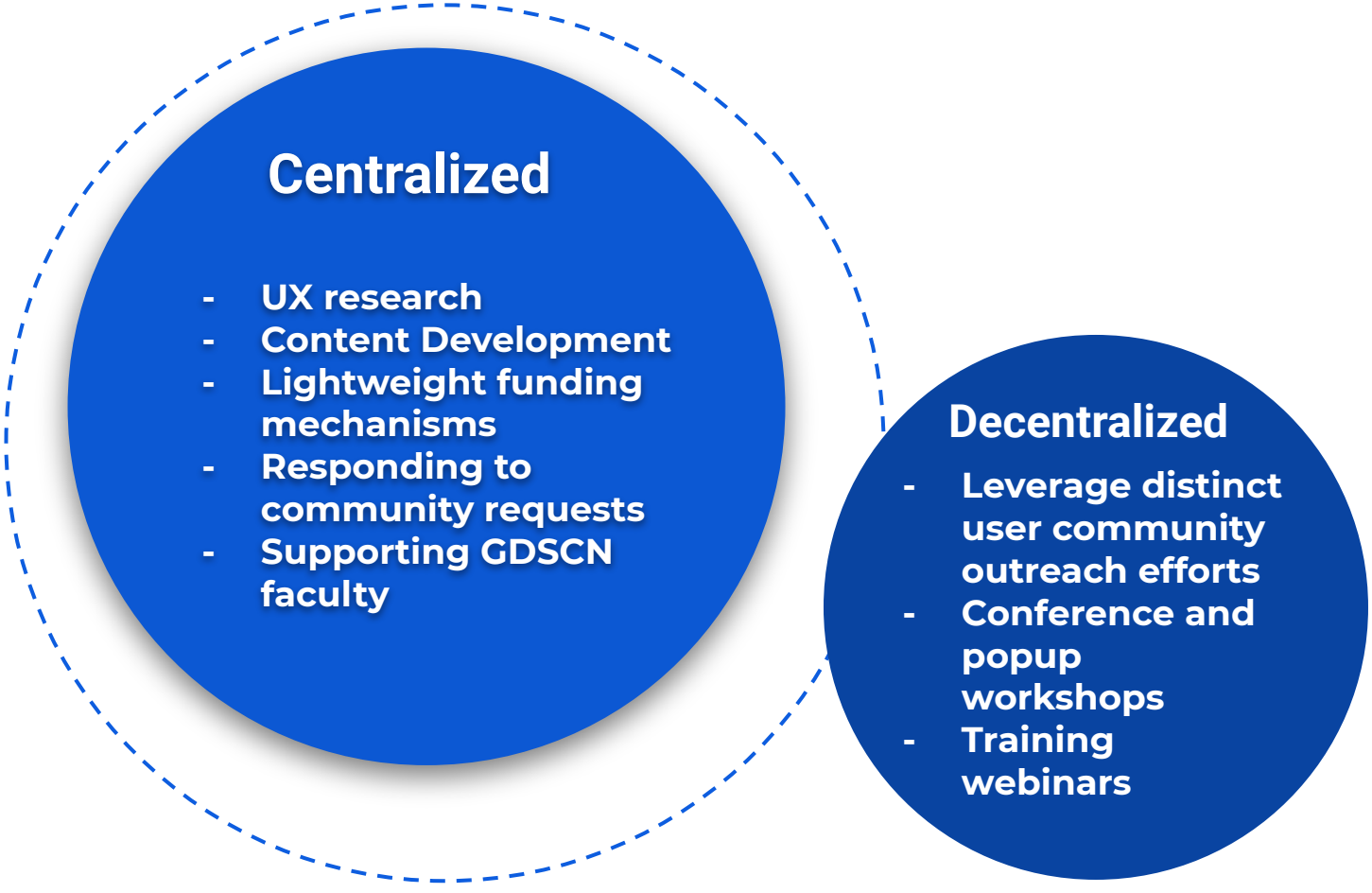
<u>Standard Computing</u>

- You buy a laptop one time
- You get that one laptop
- You pay little per use

<u>Cloud computing</u>

- You use any web browser
- You rent the computers
- You pay per hour/gigabyte/etc.

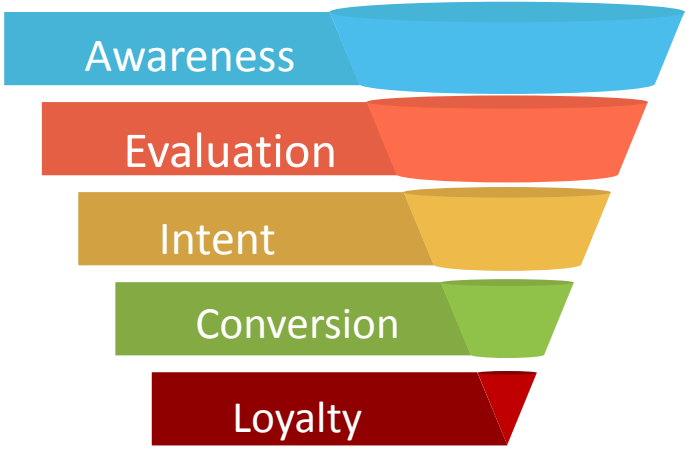
Team Design


Centralized
- UX research
- Content Development
- Lightweight funding mechanisms
- Responding to community requests
- Supporting GDSCN faculty
Decentralized
- Leverage distinct user community outreach efforts
- Conference and popup workshops
- Training webinars

Our approach

Overview of Training Activities

- Awareness and advertising
- Asynchronous Support
- Synchronous Support
- Tailored Support
- User Research



Adapted from poweredtemplate.com

Awareness

Blog Posts

Date Published	Article	Views: Publish date - Oct 26
January 22, 2021	Try out RStudio in Terra	905 (one of the top posts)
January 29, 2021	Try out Galaxy in Terra via AnVIL	332
June 9, 2021	A Galaxy of tools at your fingertips	439
June 17, 2021	How Terra fits within the AnVIL ecosystem	339
October 7, 2021	Access Terra/AnVIL resources easily with BioConductor and the AnVIL package	246

Social Media

AnVIL
@useAnVIL

Congrats to AnVIL Cloud Credits awardee @tycheleturner!

> **Tychele Turner** @tycheleturner · Oct 12
> Our @TNTurnerLab @WashUGenetics ACES manuscript is now out in Bioinformatics! doi.org/10.1093/bioinf... Code available here github.com/TNTurnerLab/AC... and it now also runs in the cloud! #genetics #genomics #bioinformatics

8:45 PM · Oct 12, 2021 · Twitter Web App

2 Retweets **5** Likes

Tweet your reply

Reply

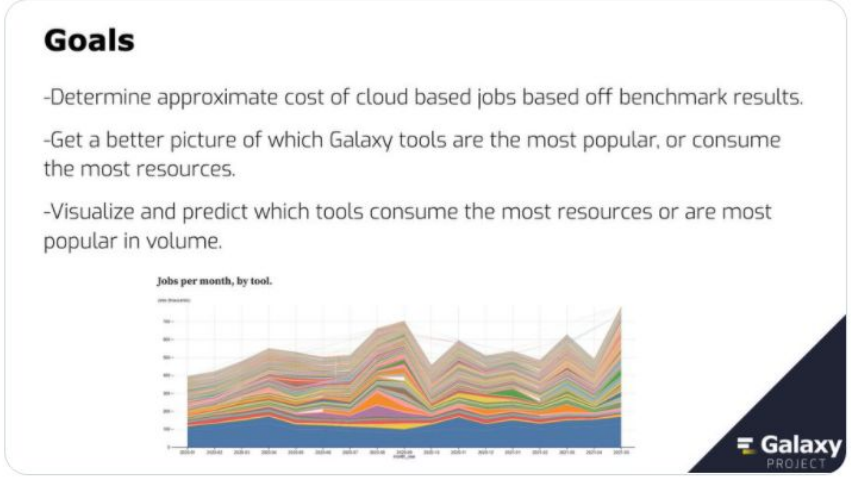
Tychele Turner @tycheleturner · Oct 13
Replying to @useAnVIL
Thank you for supporting our work!

1

AnVIL Retweeted

Galaxy Project
@galaxyproject

#UseGalaxy for @useAnVIL Cloud Cost Analysis, this Thursday at noon US Eastern galaxyproject.org/events/2021-08... Presented by @victorW65326694 and Bridget Carr. Comparing @jetstream_cloud @awscloud & @googlecloud



2:01 PM · Aug 17, 2021 · TweetDeck

7 Retweets **8** Likes

https://twitter.com/useAnVIL/

Asynchronous Support

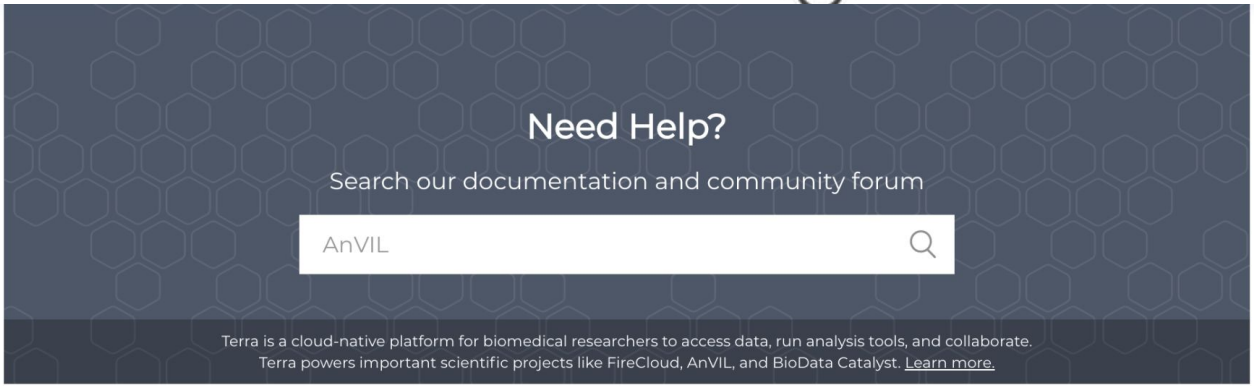
AnVIL

all categories ▸ | all tags ▸ | Latest | Top | Categories | + New Topic

Topic	Replies	Views	Activity
AnVIL Office Hours 18NOV2021 @ 11 am ET	0	11	1d
Running bwa mem on AnVIL	4	71	Sep 15
Public dissemination of derived data on AnVIL? terra	5	78	Sep 14
Welcome to Discourse	0	80	Aug 11
Error creating billing project	5	108	Jun 10
Embed images in Workspace descriptions	0	83	May 21
Load error when selecting 'Workspaces' tab terra, bioconductor, anvilpublish	7	100	May 19
Error Creating Cloud Environment	4	101	Mar 19
Workflow execution status in Terra terra	9	150	Mar 17

https://help.anvilproject.org/

AnVIL powered by Terra - Support



Terra Support > Search results

66 results for "AnVIL"

Knowledge base

AnVIL: Linking Gen3 and AnVIL on Terra
Terra Support > Documentation > Scientific Partnerships > AnVIL · Allie Hajian · 1 year ago

Next, go to your Terra Profile page - https://*anvil*.terra.bio/#profile

Launching seqr through Terra
Terra Support > Documentation > Analysis > Galaxy · Anton Kovalsky · 5 months ago

To learn more about how to use seqr and how it's implemented in *AnVIL*

Community

Workspace email notifications (e.g. AnVIL) 0
Terra Support > Community > Known Issues · Tiffany Miller · 1 year ago

Many people are getting email notifications with the subject "Terra: Data has been added to *AnVIL*...workspace" and would

Add the "AnVIL" package in the cloud environments by default 1
Terra Support > Community > Feature & Documentation Requests · Eugenio Mattei · 5 months ago

The *AnVIL* package is an essential tool to interface with the

Channels for support:

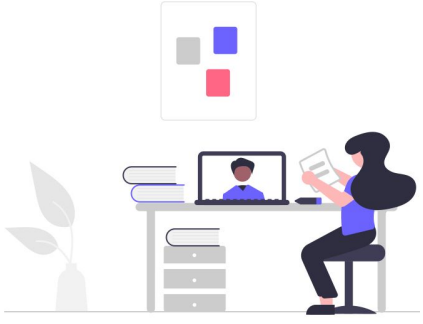
- AnVIL Discourse (new this year!)
- Terra Support forum
- Contact Us button via anvil.terra.bio
- Email (for Terra questions)

https://support.terra.bio/hc/en-us

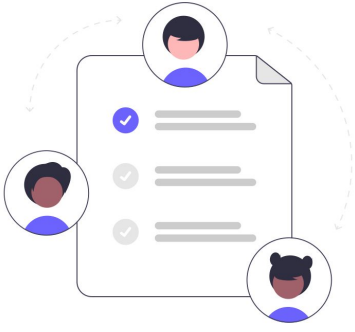
Persona-oriented approach for training materials

PIs

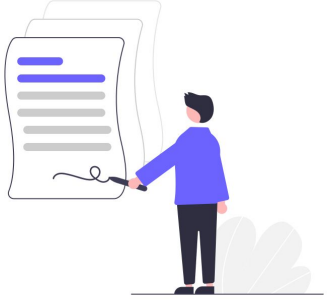
Analysts



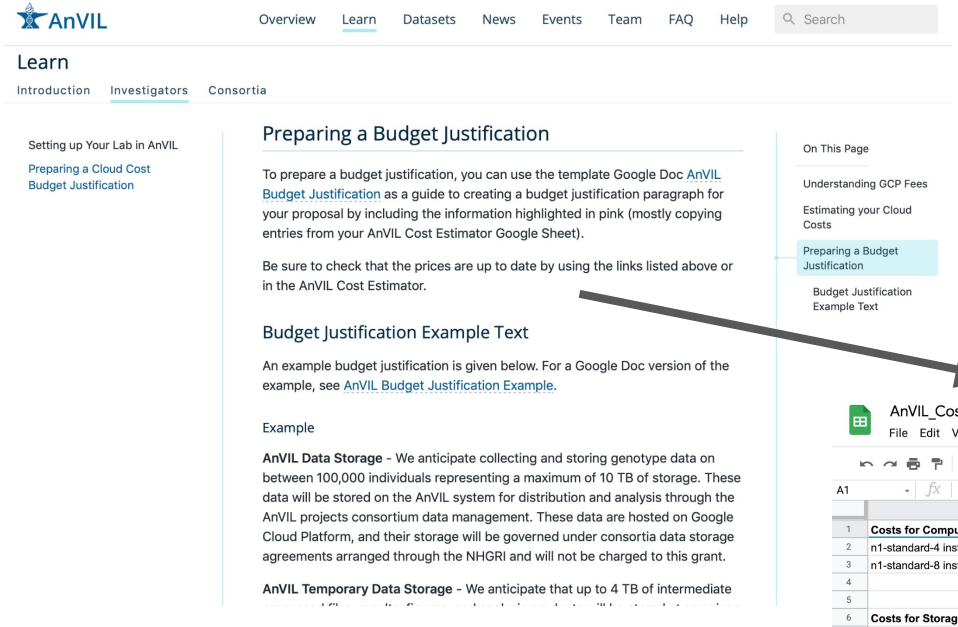
Consortia



Teachers



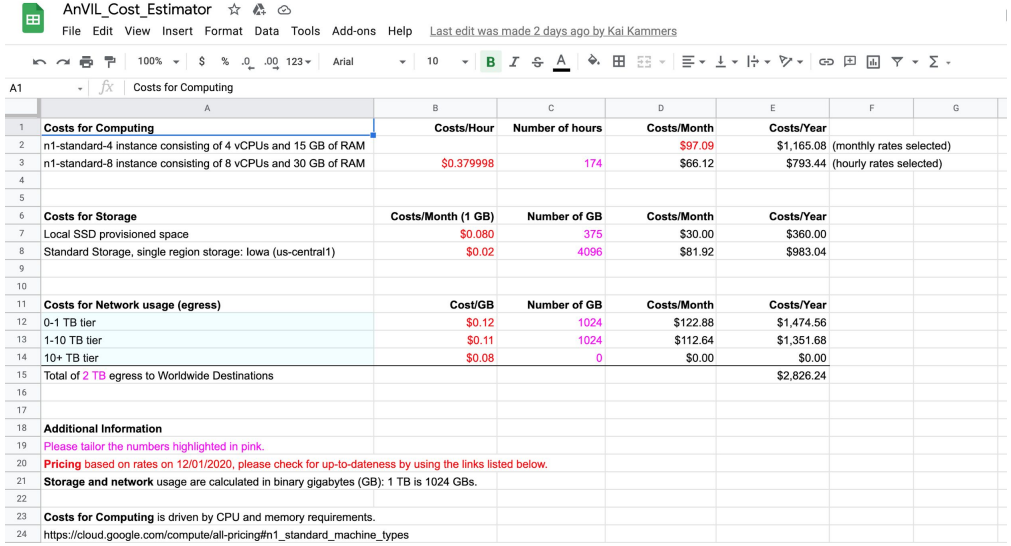
Budget Templates for AnVIL



Completed: AnVIL Getting Started Guide - 1.0

Upcoming:

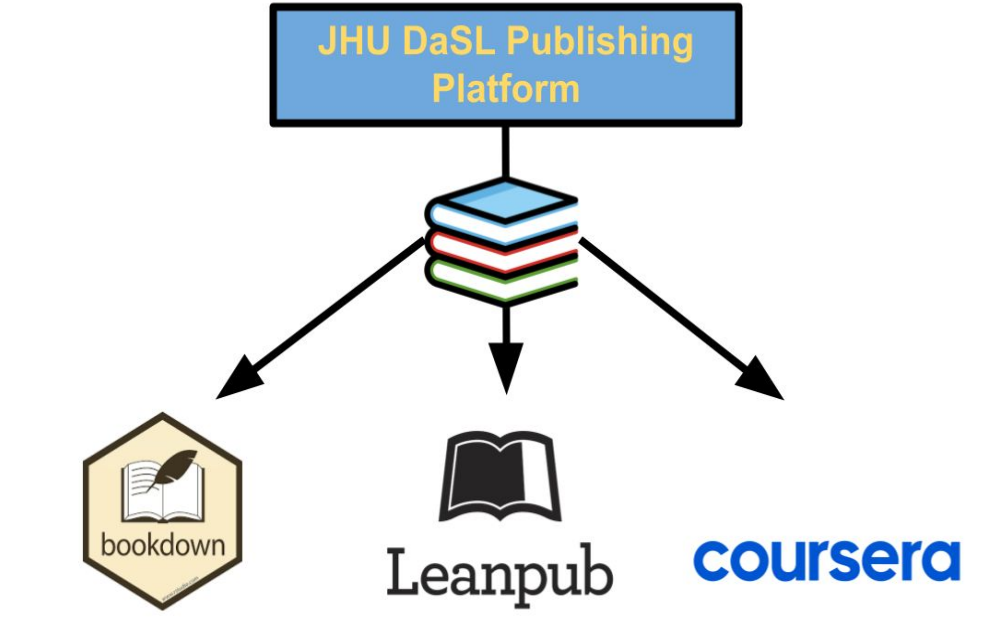
Teacher Guide, Exercises Book & Galaxy Guide



	A	B	C	D	E	F	G
1	**Costs for Computing**	**Costs/Hour**	**Number of hours**	**Costs/Month**	**Costs/Year**		
2	n1-standard-4 instance consisting of 4 vCPUs and 15 GB of RAM			$97.09	$1,165.08	(monthly rates selected)	
3	n1-standard-8 instance consisting of 8 vCPUs and 30 GB of RAM	$0.379998	174	$66.12	$793.44	(hourly rates selected)	
4							
5							
6	**Costs for Storage**	**Costs/Month (1 GB)**	**Number of GB**	**Costs/Month**	**Costs/Year**		
7	Local SSD provisioned space	$0.080	375	$30.00	$360.00		
8	Standard Storage, single region storage: Iowa (us-central1)	$0.02	4096	$81.92	$983.04		
9							
10							
11	**Costs for Network usage (egress)**	**Cost/GB**	**Number of GB**	**Costs/Month**	**Costs/Year**		
12	0-1 TB tier	$0.12	1024	$122.88	$1,474.56		
13	1-10 TB tier	$0.11	1024	$112.64	$1,351.68		
14	10+ TB tier	$0.08	0	$0.00	$0.00		
15	Total of 2 TB egress to Worldwide Destinations				$2,826.24		
16							
17							
18	**Additional Information**						
19	Please tailor the numbers highlighted in pink.						
20	**Pricing** based on rates on 12/01/2020, please check for up-to-dateness by using the links listed below.						
21	**Storage and network** usage are calculated in binary gigabytes (GB): 1 TB is 1024 GBs.						
22							
23	**Costs for Computing** is driven by CPU and memory requirements.						
24	https://cloud.google.com/compute/all-pricing#n1_standard_machine_types						

AnVIL Content

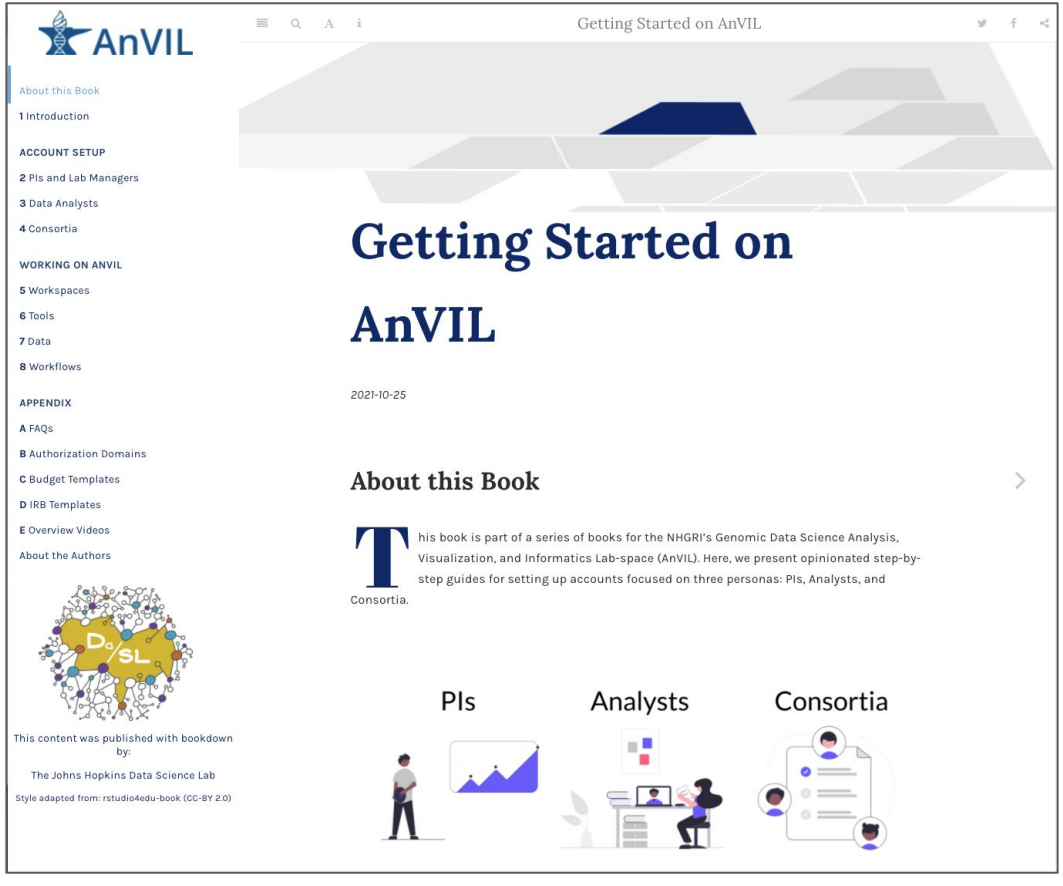
JHU DaSL EdTech Stack



Free for learners	Free or fee for learners	Membership fee for learners
No certification	✅ Includes certification of completion	✅ Includes certification of completion

Publish and maintain on all 3 platforms!

AnVIL Getting Started Guide - 1.0



https://jhudatascience.org/AnVIL_Book_Getting_Started/

Opinionated guide for PIs/Lab Managers, Analysts, and Consortia Members

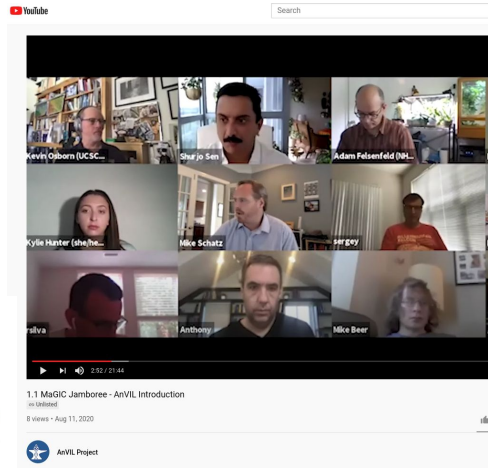
Walks through:

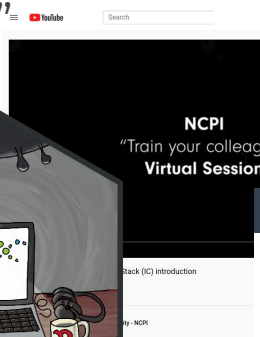
- Billing Account Setup
- How to work on AnVIL
  - Workspaces
  - Tools
  - Data
  - Workflows
- Budget and IRB templates

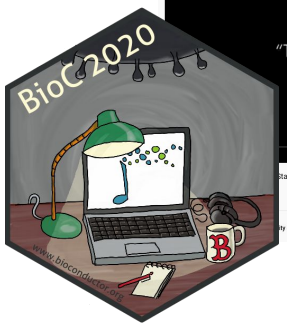
Synchronous Support

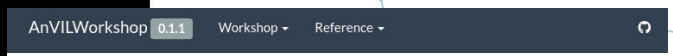
Synchronous Events

- Galaxy Webinars
- Health Disparities Researchers
- ASHG
- T2T/HPRC
- Bioconductor Community Conference
- Bioinformatics Community Conference
- eMERGE Researchers
- ISMB
- MaGIC Jamboree
- NCPI “Train your colleague”
- PAG
- GDSCN
- + many more









Introduction to the Terra/AnVIL Cloud-based Genomics Platform

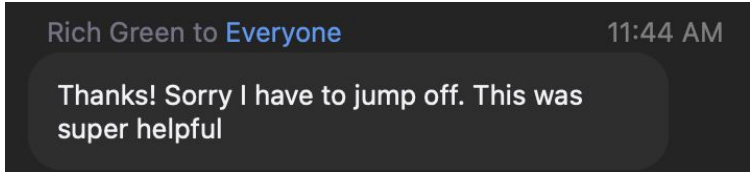
Galaxy as an educational tool and community resources for undergraduate training

PAG 2020

Office Hours

Attendees:

- Data Coordinating Center member
- Data analyst applying ML techniques
- AC2 awardee

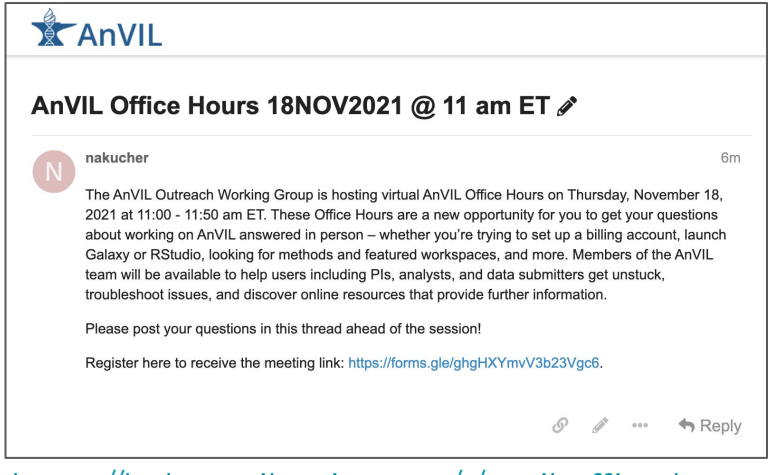


Discussed solutions to roadblocks with:

1. Set up/Billing
2. Data access
3. Analysis
4. Future work and support needed

Highlighted **help.anvilproject.org** as a resource to get help for new issues and questions in the future.

Next Office Hours: 18NOV



https://help.anvilproject.org/t/anvil-office-hours-18nov2021-11-am-et/56

Tailored Support

AC2 Awardees

Alex Greiner | The University of Iowa | Graduate Student
- "Burden analysis of inherited cardiac arrhythmia genes in epilepsy"

Melissa Suzanne Cline | UC Santa Cruz Genomics Institute | Principal Investigator
- "Leveraging AnVIL and Terra for secure collaboration on genetic variant interpretation"

Andrew Davidson | University of California | Graduate Student
- "Comprehensive characterization of transposable element expression across human tissues"

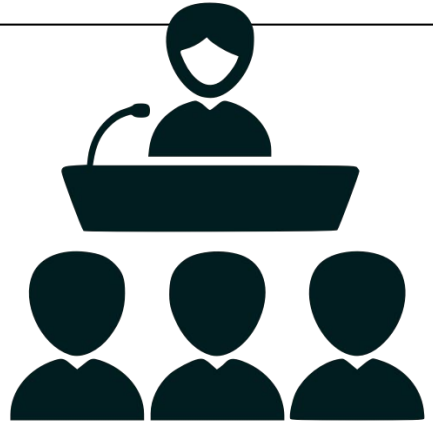
Anahita Khojandi | University of Tennessee-Knoxville | Associate Professor
- "Deep Learning for Accurate Tissue-Specific Prediction of Gene Expression in Large Deeply-Phenotyped Population"

Anshul Kundaje | Stanford University | Principal Investigator
- "Deciphering cis-regulatory syntax of a transcription factor binding atlas with interpretable deep learning models"

Tychele N. Turner | Washington University in St. Louis | Principal Investigator
- "A k-mer based approach to assess copy number in PacBio HiFi data"

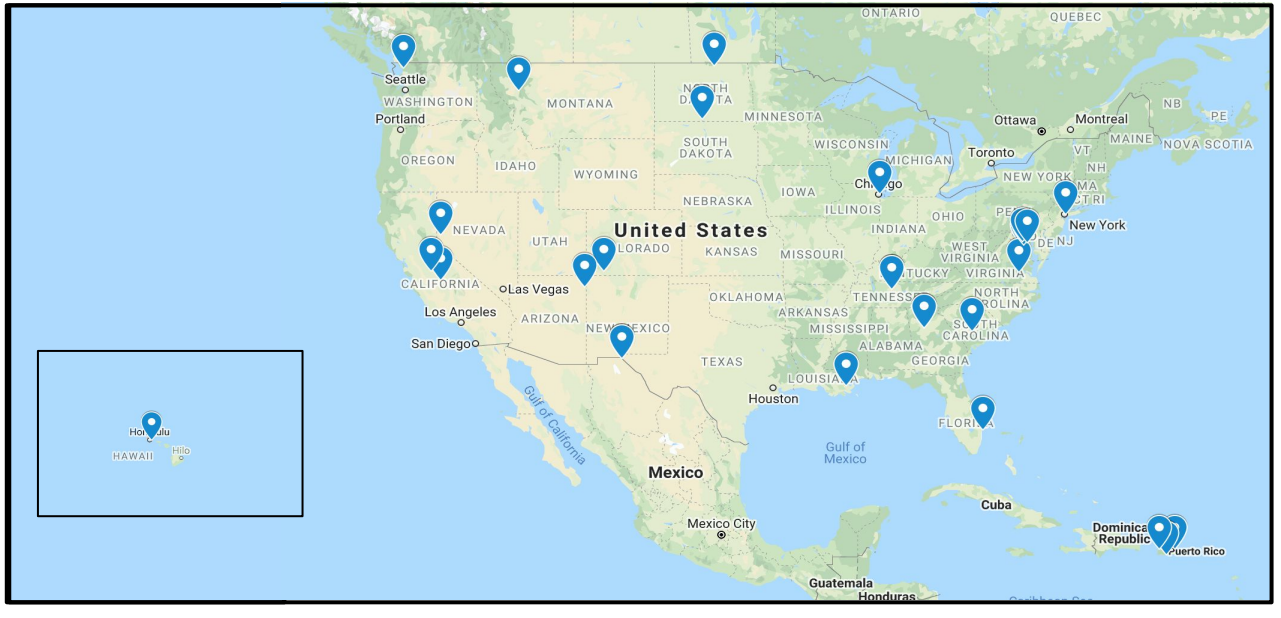
Meet and Greets

Hosted 2 sessions for awardees to meet, present their work and progress, and share their experiences using AnVIL



Genomic Data Science Community Network (GDSCN)

To broaden the spectrum of diverse institutions active in bioinformatics and genomic data science, we seek partnerships with educators and researchers at Historically Black Colleges and Universities (HBCUs), Minority Serving Institutions (MSIs), Tribal Colleges and Universities (TCUs), and Community Colleges (CCs).



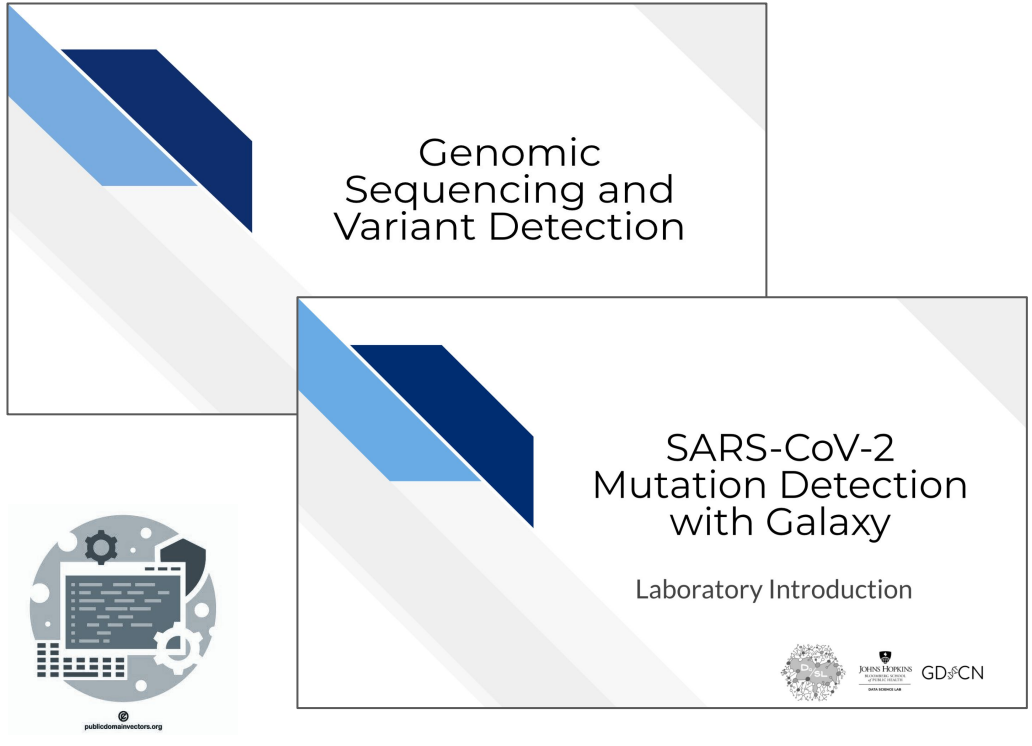
http://www.gdscn.org/

GDSCN Activities

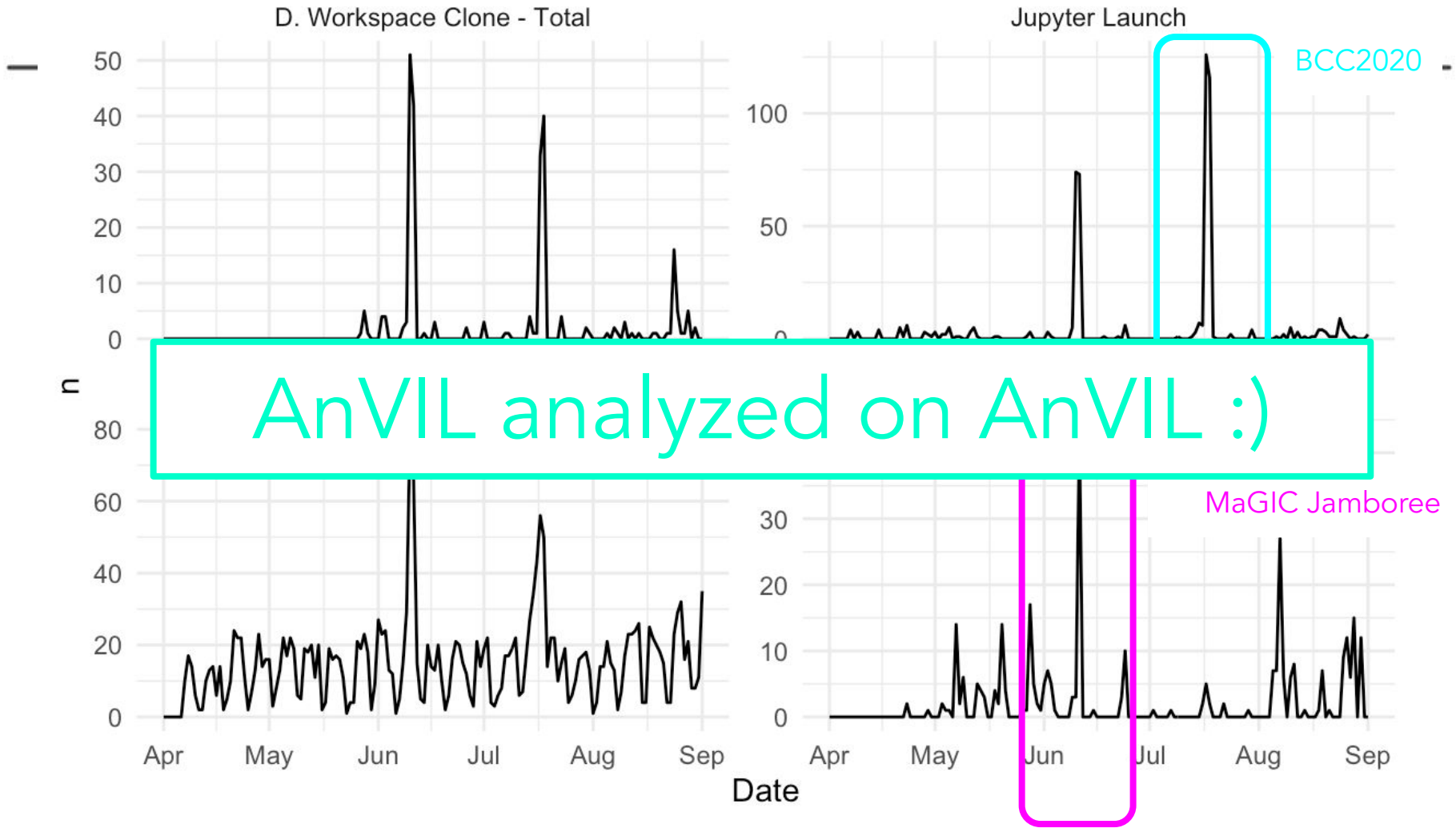
Events

1. **Kickoff Meeting**
   March 2021
2. **Symposium #2**
   April 2021
3. **Symposium #3**
   November 2021

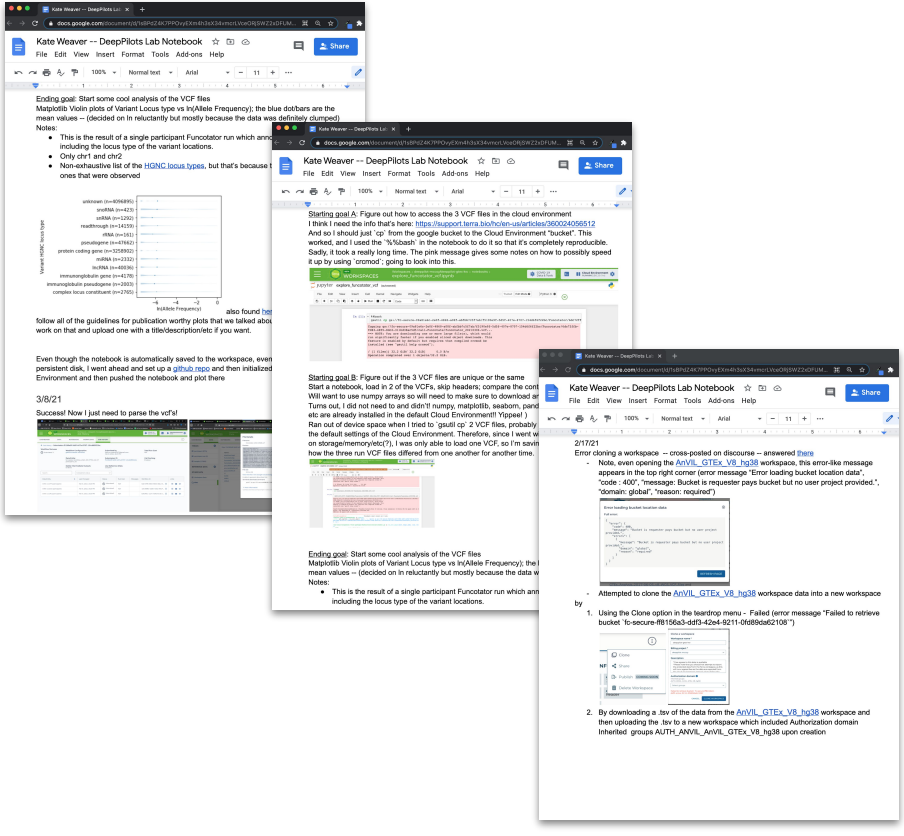
Developing Lectures and Exercises



User Research



DeepPilots - harnessing researcher experience



By tracking the progress of researchers performing novel research on the cloud we can:

- Identify our largest roadblocks
- Chart a course to circumvent these roadblocks
- Build appropriate tutorials and help documents to streamline transition to research on AnVIL

Goal

To build a happy, big, diverse, vibrant AnVIL user community doing teaching & research.

Vision for the Future

We built it, now we want tons of users to come to AnVIL!

1. Scale up our asynchronous (scalable) and synchronous (deep) training efforts
2. Continue to develop content and tools that can support platform changes
3. Build the AnVIL community (GDSCN, clinical, AnVIL Discourse)
4. Leverage our user research to improve user experience
    a. *Simplified/scalable financial support through a free tier*
    b. Easier access points to tools through the website
    c. Simplified billing/data management/user management interfaces