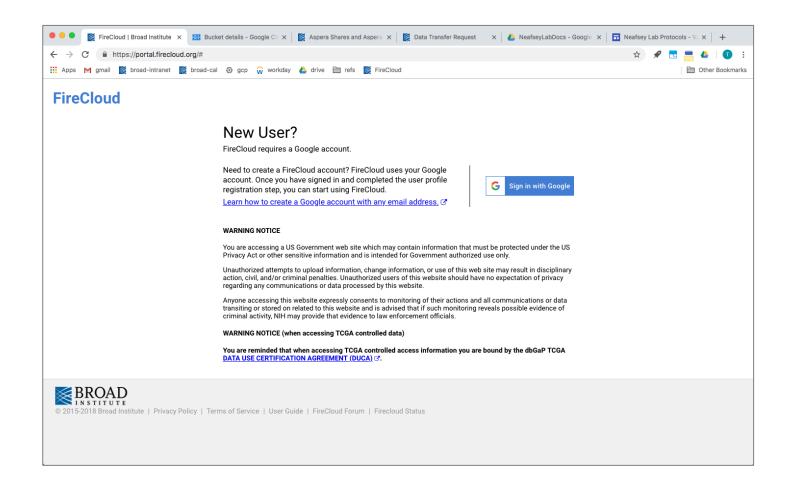
FireCloud Sign-in

Go to <u>portal.firecloud.org</u>

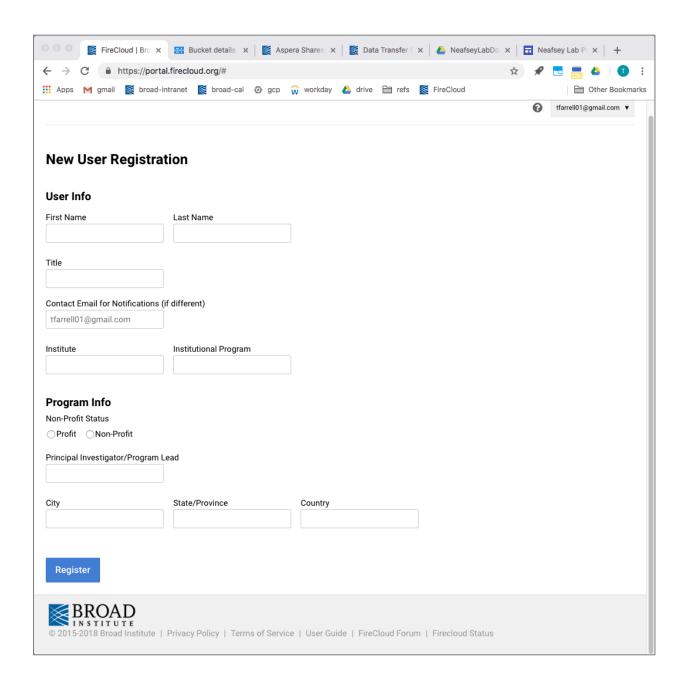
- Click Sign in with Google
 - (use your Broad Institute email over a personal Google account)



FireCloud Registration

 If you are a new user, you will be directed to the registration page

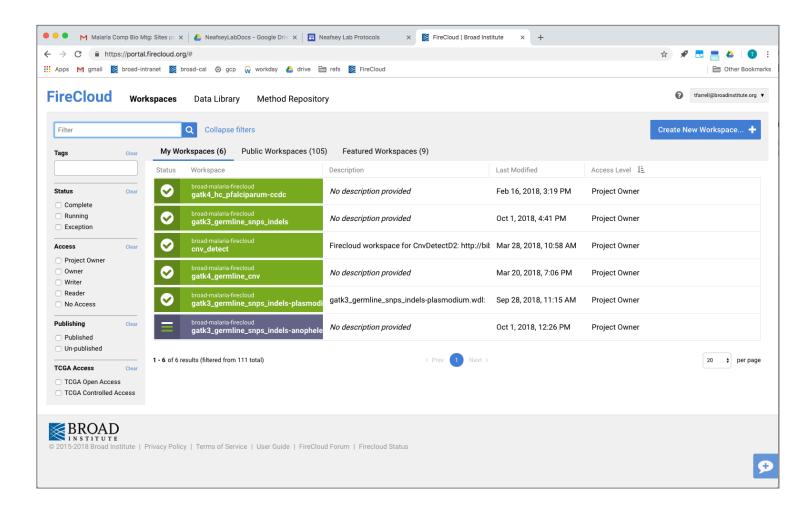
Fill out your information accordingly



FireCloud Workspaces

 Once signed-in/ registered, you should be directed your workspaces page

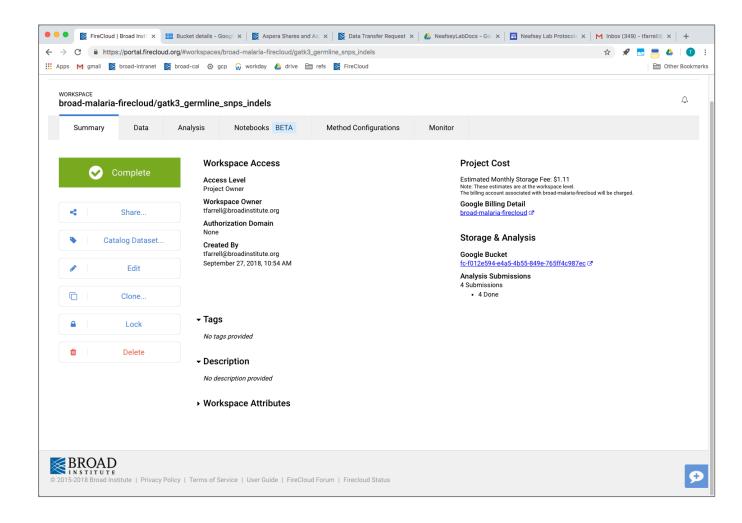
 Workspaces organize methods, data models (when applicable) and job monitoring in one place



FireCloud Workspaces

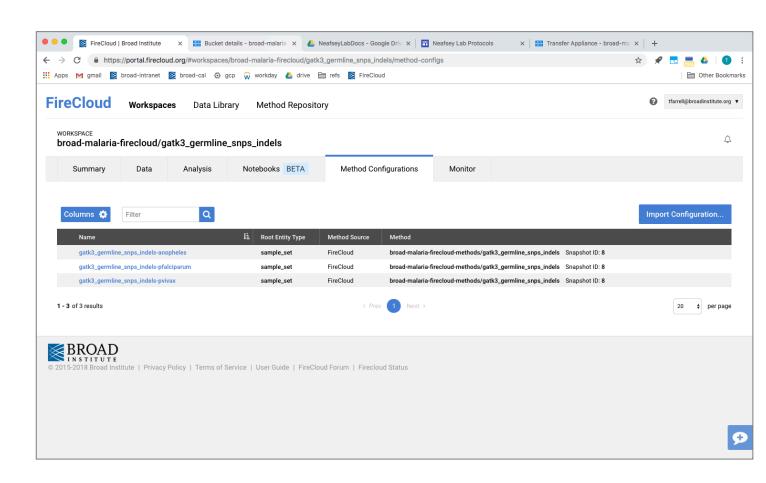
 When you select a workspace, the summary page looks like this

 The data models, methods and job monitoring can be found in the Data, Method Configurations and Monitor tabs respectively

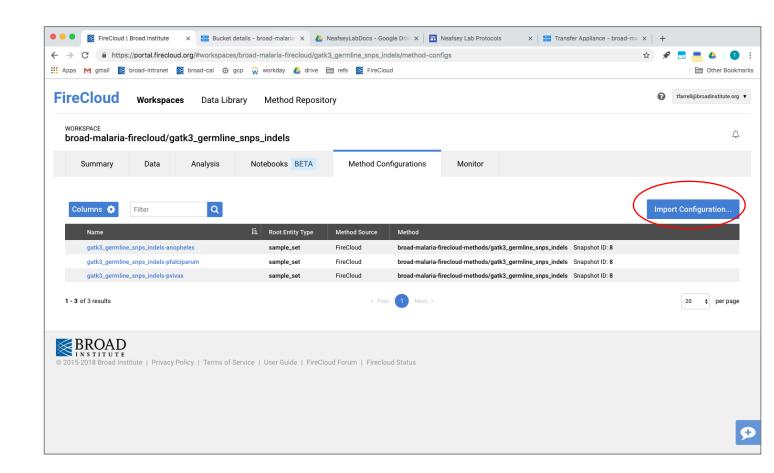


broad-malaria-firecloud/gatk3_germline_snps_indels

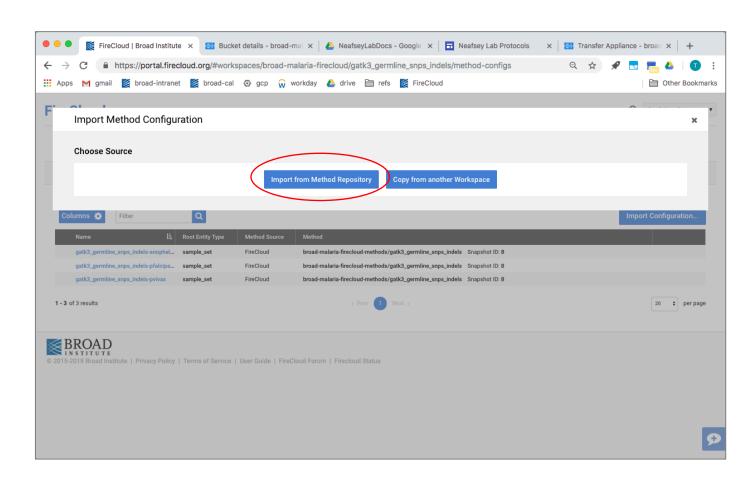
- The main workspace for malaria variant-calling is gatk3_germline_snps_indels
- If you select this workspace, and then click the Method Configurations tab, you will see 3 organism-specific template configurations:
 - anopheles : A. gambiae
 - pfalciparum : P. falciparum
 - pvivax : P. vivax



- When starting a new variant-calling analysis, you'll want to create a new configuration from one of the 3 templates, depending on which organism your data was derived from
- To do this, go to the workspace gatk3_germline_snps_indels and the Method Configurations tab
- Then click on **Import Configuration...**

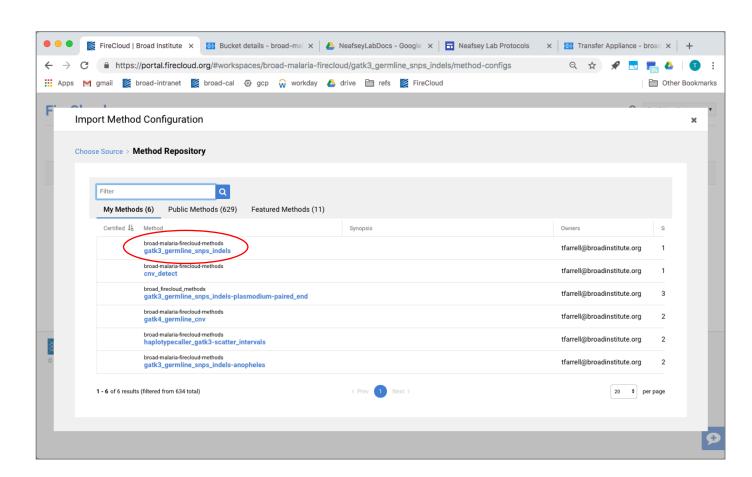


 Then click on Import from Method Repository



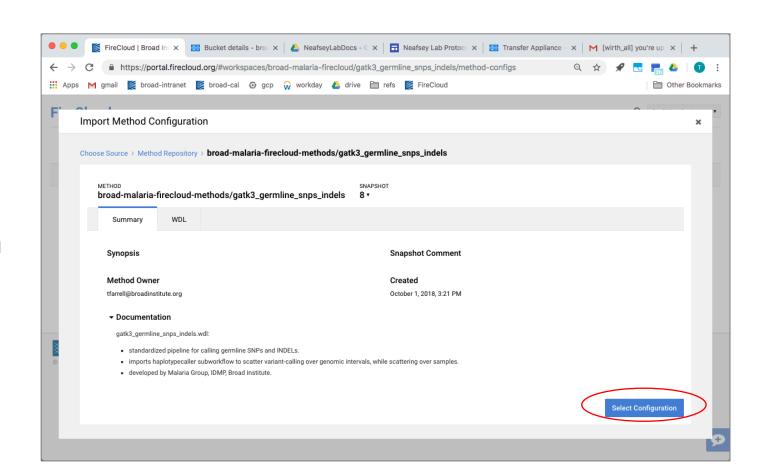
 You will see a list of methods in your Methods Repository

Select the method gatk3_germline_snps_indels



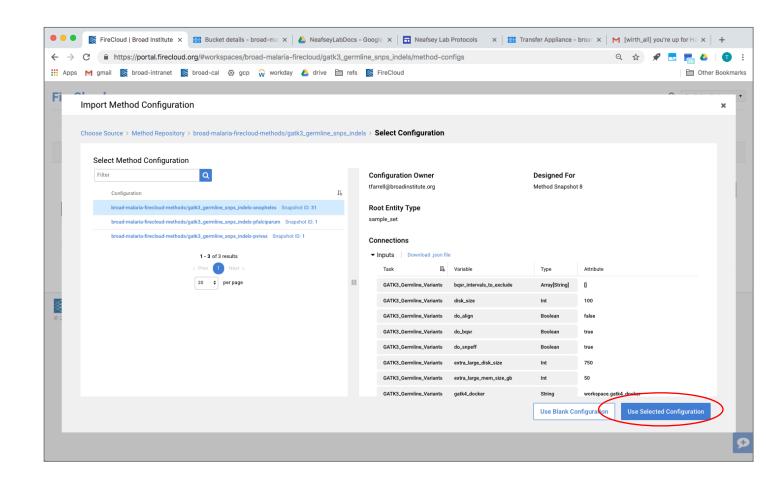
You will see a summary of this method

 Then click Select Configuration to select one of the template configurations



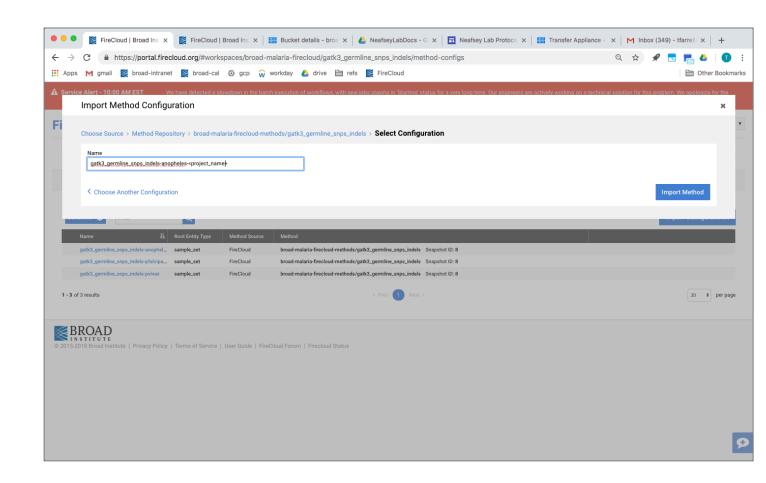
 You will see the list of template configs in the left panel. If you click on any of these, you will preview its values in the right panel

 Select the template config specific to your organism, and then click Use Selected Configuration

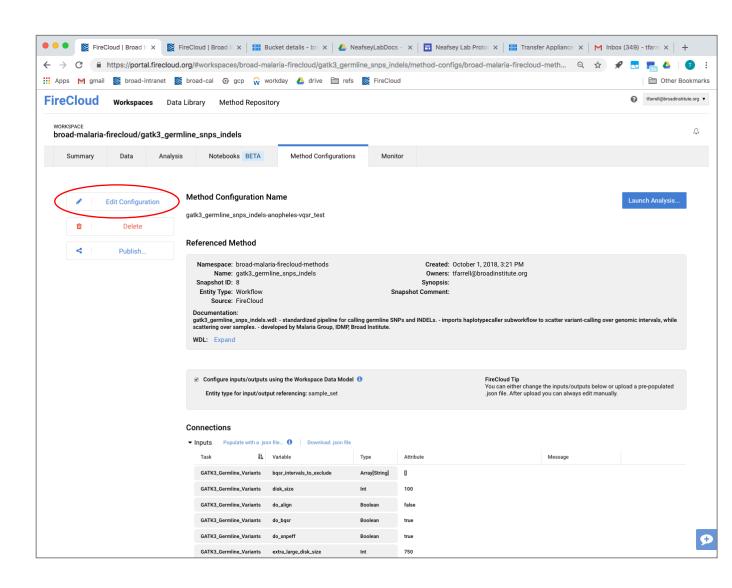


- You will then need to rename this config to make it specific to your project
 - It's recommended that you follow this convention: gatk3_germline_snps_indels-<organism>-<project>

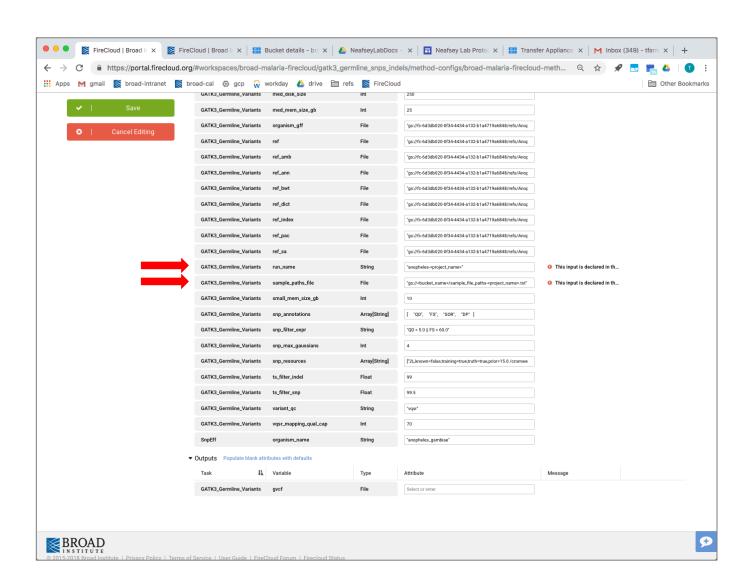
Then click Import Method



- This will create a new method config specific for your project, which will serve as a record of the config values for this run
- You can then change any of the config values as your analysis might require
- To do this, click Edit
 Configuration

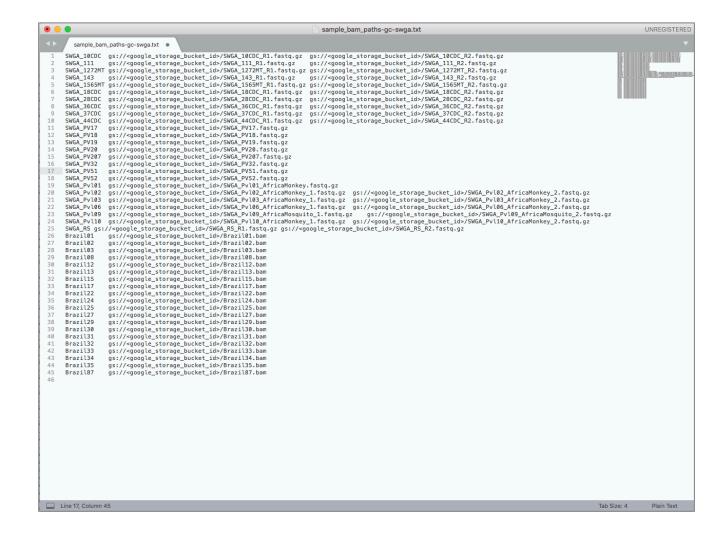


- If you don't wish to change the config values from the defaults (which follow GATK best practices), the only values that require input are run_name and sample_paths_file
- For run_name, this will served as the prefix for the final GVCF:
 <run_name>.snpeff.g.vcf.gz
- For sample_paths_file, this is the Google Storage path to the tabdelimited file listing your samples and their corresponding raw data file



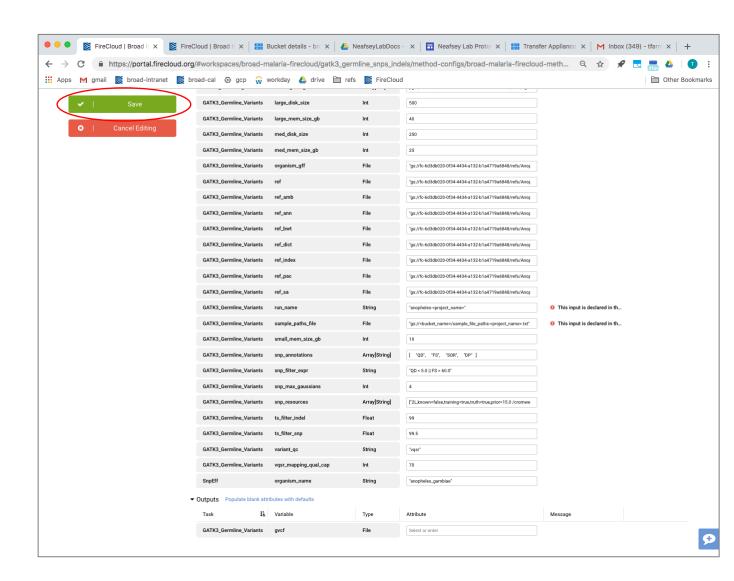
 Each line in your sample_paths_file will need to follow either of the following formats:

```
<sample_id> gs://<sample_id>.[bam|fastq.gz]
OR
<sample_id> gs://<sample_id>.R1.fastq.gz
gs://<sample_id>.R2.fastq.gz
```

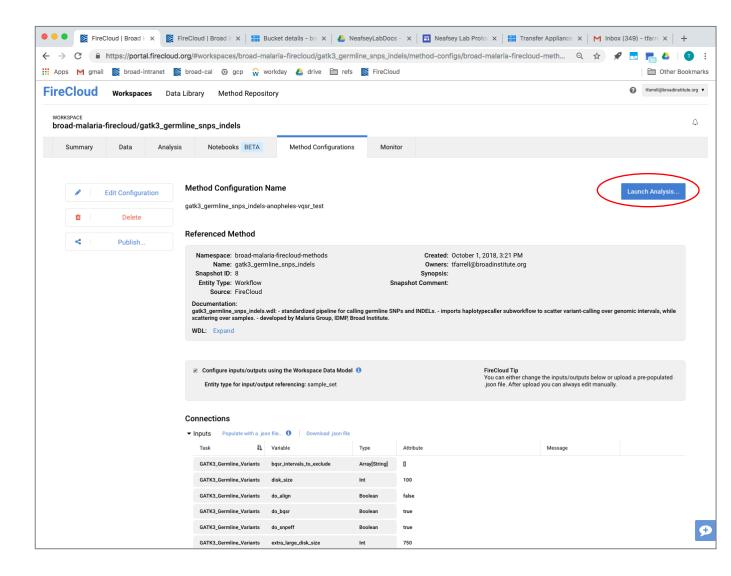


 Once you've filled in your run_name and sample_paths_file, click the Save button

 After this step, you'll want to upload your data and then you can launch your analysis



broad-malaria-firecloud: Uploading data



broad-malaria-firecloud: Launching an analysis

Once your configuration is complete

 This will save this configuration for future. And now that it's complete, you can launch the analysis

