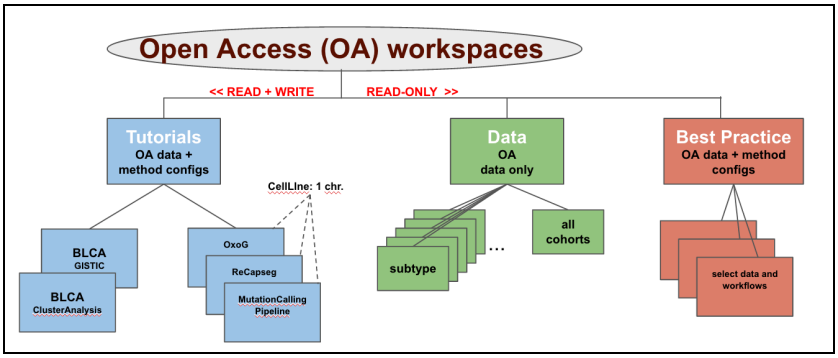
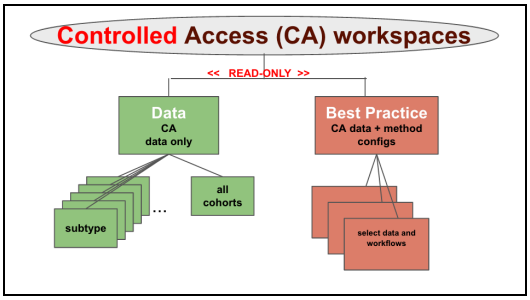
Pre-populated TCGA Workspaces

FireCloud includes three types of pre-populated workspaces, **Workshop** (data and basic workflows), **Data** (data-only) and **Best Practice** (data and workflows). These contain either TCGA Open Access or Controlled Access data. Workspaces containing TCGA Open Access data will be available to the public. TCGA Controlled Access data will require dbGaP authorization (see below).



Workshop workspaces contain basic workflows that demonstrate FireCloud functionality. In order to clone and use these workspaces, you must be granted access to or set up a FireCloud Billing Project.



Data workspaces contain curated data sets, either Open Access or Controlled Access data. Because TCGA data contains many replicates of files, for example, failures and re-aggregations, multiple tissue types (e.g., frozen, FFPE, and blood), and numerous aliquots, it is challenging to select the best data set for an analysis. To simplify this process we have loaded the **Data** workspaces with FireCloud-curated 'best' datasets based on our knowledge and experience using TCGA data. In order to clone and run analyses on Data workspaces, you must be granted access to or set up a FireCloud Billing Project. You can import methods or additional data into the cloned workspace, as desired. Access to Controlled Access TCGA data will be granted to users with the appropriate authorization (see below).

Best Practice workspaces contain either Open Access or Controlled Access data and select Broad Institute best practice workflows. These workspaces will be loaded with curated data sets, attributes, best practice workflows and results. In order to clone and run analyses using Best Practice workspaces, you must be granted access to or set up a FireCloud Billing Project. You can import methods or additional data into the cloned workspace, as desired. Access to Controlled Access TCGA data will be granted to users with the appropriate authorization (see below).

Accessing TCGA data in FireCloud

TCGA Open Access data is available to all FireCloud users. Open access data will be found in the pre-loaded workspace.

TCGA Controlled Access data is accessible to users who have dbGaP authorization to use Controlled Access data.

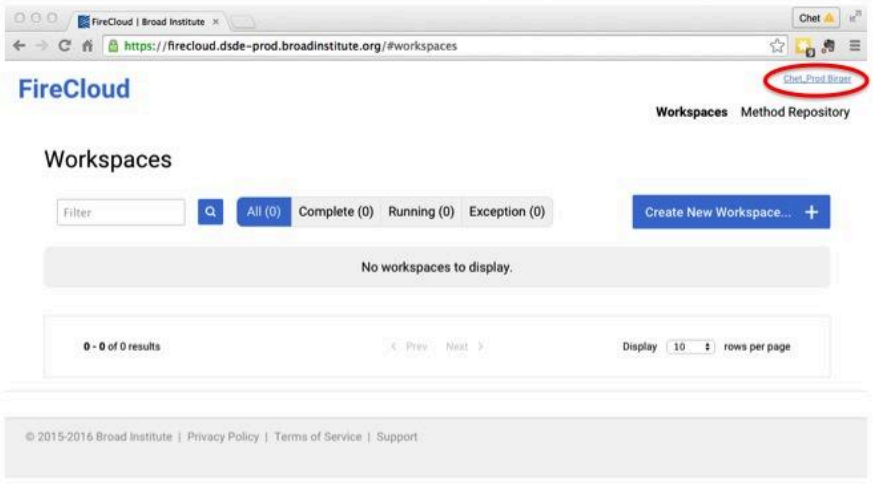
To access workspaces in FireCloud containing Controlled Access data, you must

1) have an eRA Commons or NIH account with dbGaP authorization,
2) link your FireCloud account to that eRA Commons or NIH account.

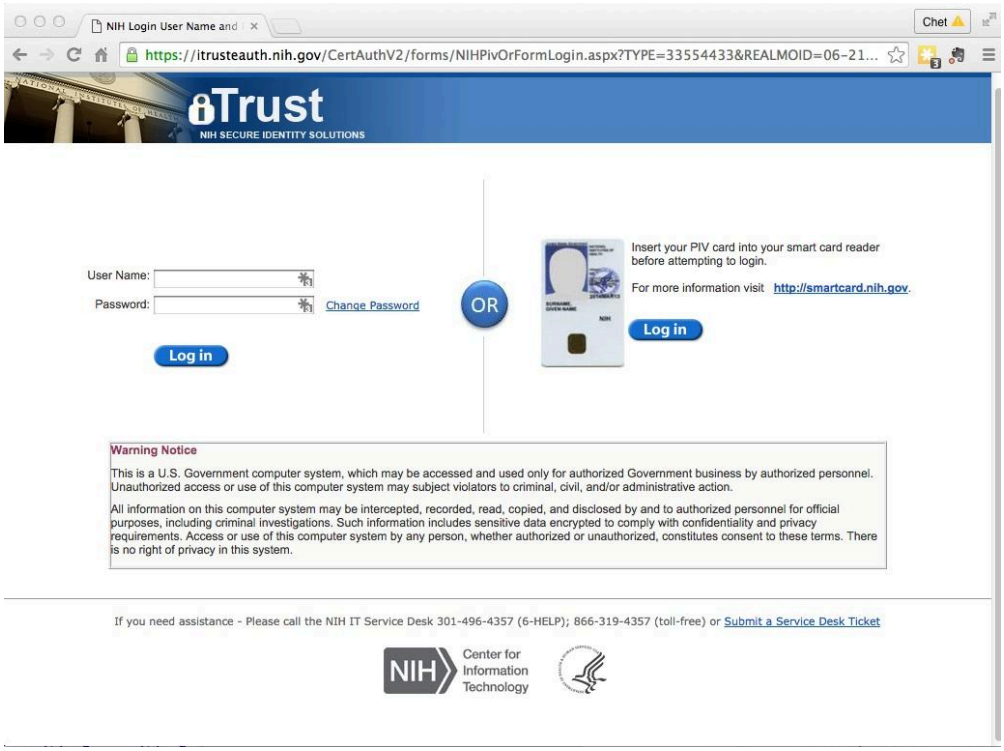
Instructions

Instructions to set up an eRA Commons account and gain dbGaP authorized access can be found here.

To link your eRA Commons / NIH account, go to the User Profile page by clicking on your name displayed in the upper right-hand corner of the FireCloud web interface.



A link will appear at the bottom of the FireCloud User Profile page labeled “Log-in to NIH to link your account”. After clicking on this link, you will be redirected to NIH’s federated identity service for authentication of eRA Commons and NIH identities.



You should log in with your eRA Commons / NIH username and password on this page. If authentication is successful, FireCloud will create the linkage between your FireCloud account and your authenticated NIH / eRA Commons identity.

FireCloud will then check whether the NIH / eRA Commons identity is dbGaP authorized. If you are dbGaP authorized, you will be given access to TCGA Controlled Access workspaces.

For security purposes, this linkage will expire within a set amount of time. FireCloud requires users with linked accounts to refresh their linkage by reauthenticating their NIH / eRA Commons identities. In addition, FireCloud removes users’ access to TCGA controlled workspaces if they lose their dbGaP authorization.