Service Workbench User Guide

Table of Contents

Account Structure....4
Local Users....4
Sidebar Introduction....4
Sidebar: Researcher's view....5
Dashboard....5
Studies: Introduction....6
Creating a Study....6
Studies Page....7
Changing organization-categorized Study permissions....7
Creating a Study and uploading files....7
Sharing a Study....8
Data Sources: Introduction....8
Registering an external Study with Service Workbench....8
Onboarding an external Study account....9
Removing External Studies....10
Workspaces Introduction....11
Creating a new Workspace....11
Accessing a Workspace....12
Connect to SageMaker and EMR workspace....12
Connect to EC2 Linux....13
Connect to EC2 Windows....13
Connect to RStudio....14
Start and Stop workspace....14
Common connection issues....15
Terminating a Workspace....15
Administrator's view....16
Dashboard....16
Auth Introduction....16

Users Introduction....16
User Roles....16
Adding Federated Users....17
Add a single federated user....17
Add multiple federated users....17
Accounts....17
Projects Introduction....17
Create a New Project....17
Adding a User to a Project....18
Indexes Introduction....18
Create a New Index....18
AWS Accounts....18
Introduction to AWS Accounts....18
Create an AWS Account....19
Set AWS Account Budget....20
Roles and Permissions....20
Master Account Role....20
Master Role Permissions....21
Cross Account Execution Role....21
Permissions....22
Workspaces: Introduction....**Error! Bookmark not defined.**
Studies: Introduction....23
Creating a Study....25
Studies Page....25
Sharing a Study....26
Data Sources....27
Introduction....27
Registering an external Study with Service Workbench....27
Onboarding an external Study account....28
Removing External Studies....29
Workspaces Introduction....**Error! Bookmark not defined.**
Creating a new Workspace....**Error! Bookmark not defined.**
Terminating a Workspace....**Error! Bookmark not defined.**

Best Practices 29
Multiple Deployment Environments 29
Amazon Inspector 29
AWS CloudTrail 30
AWS Shield 30
CI/CD 30
How the CI/CD pipeline works 30
Deploying the CI/CD Pipeline 32

Account Structure

Service Catalog uses three kinds of accounts, whose names are used in this guide. *Master*and *Member* accounts are terms referring to AWS Organizations, while *Main* account is a Service Catalog term.

- **Main**: The account within which Service Catalog is deployed. The main account will be billed for all AWS usage charges for the Service Catalog deployment itself: the S3 bucket holding the website, and other resources not created by a user. Service Catalog may be deployed in a master account (the account holding the Organization), or a member account (within the Organization). In either case, this account is called the Service Catalog *Main* account.
- **Master**: The account hosting the AWS Organization. The master account is responsible for the billing of the member accounts within the Organization.
- **Member**: An account within an AWS Organization. When you create an account using Create an AWS Account, that account is created as a member of the Organization.

See also in the source code:

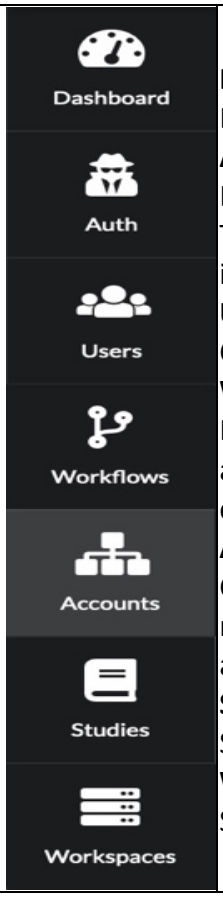
- README.md
- main/solution/prepare-master-acc/README.md

Local Users

Local users are created only within the solution and their credentials stored in DynamoDB. This is a fast way to get an installation working since the alternative is to integrate with an AD.

Sidebar Introduction

This section of the documentation aims to provide an overview of how to operate the web-based user portal that is provided as part of this solution. The User portal presents itself with a menu bar on the left side (sidebar) of the screen with a set of icons linking to different functionality of the web interface.
Based on the role a user holds when logged into the user interface, some or all menu items will be available. The goal is to present a user only with those entries that are relevant for their work and not confuse them with additional, irrelevant, entries.
The following entries are available in the sidebar based on a user's role.



Dashboard
Displays spending over the past 30 days
Auth
Lists the Authentication Providers configured in Service Workbench. This includes Default (local accounts) and any Active Directory id entity providers.
Users
Create and manage local and federated users and roles.
Workflows
Monitor and see the history of operations initiated by the administrative interface. This includes provisioning an Account and creating and deleting a Workspace
Accounts
Create and manage accounts. See: Account Structure for the relationship between AWS accounts and Service Workbench accounts, and Account for administrative actions.
Studies
See: Studies
Workspaces
See: Workspaces

Sidebar: Researcher's view

Dashboard

The dashboard is visible to all users that log in to the solution. It is also the first screen visible to a user after logging into the web interface. Based on the role a user holds they can see more or less data inside the dashboard's widgets.

The dashboard displays:

- The **Index** costs for the past 30 days. This is the cost, a certain index or cost center has allotted over the course of the last rolling 30 days. This includes all workspaces started for this index, regardless of the user who started or owns them.
- The **Workspaces** costs for the past 30 days. This represents the workspaces owned by the specific, logged in, user over the course of the last 30 days.
- Yesterday's **Workspaces** costs.
- A breakdown of the **Index** costs for the past 30 days per user.

Studies: Introduction

Service Workbench provides multiple ways to connect compute workspaces to data (Studies) saved in Amazon S3. As a researcher, you can also use it to search data using tools specific to your requirements.

Using Service Workbench, an organization can share Studies with other organizations with access controls.

There are three types of Studies available in Service Workbench. You can mount multiple Studies, of any type, to your Workspaces.

Type	**Description**
My Studies	Specifies Studies created by users. Use this option to work on datasets that are exclusive to you or that are used specifically for your research.
Organization Studies	Specifies Studies that can be shared with other users. It contains data that had been collected by an organization or is licensed. You can grant or deny users access to this data in order to comply with regulations or licensing restrictions on the data.
Open Data	Service Workbench provides access to Open Data on AWS data by frequently scanning the set of AWS open datasets and adding new datasets to this category. This can include 1,000 genomes and other datasets openly available through Amazon.

Service Workbench can host data for My Studies and Organization Studies internally, in the **studydata** S3 bucket created in the AWS account where Service Workbench was deployed. The application can also provide access to Studies hosted in external S3 buckets in other AWS accounts using the *Datasets* page.

Creating a Study

Studies are data sets. Admin, Internal Researchers and External Researchers can search for Studies to perform research on. They can also create new Studies and upload data to Studies.

Restrictions: I the disableAdminBYOBSelfAssignment flag is enabled, an administrator will be able to navigate to the data sources and register a study, but will only be able to assign a researcher as the study admin.

To create a new Study, follow these steps:

1. In the portal navigate to the **Studies** page using the menu on the left.

2. Click the **Create Study** button.
3. Provide an ID for the Study in the **ID** field.
4. Choose either **My Study** or **Organizational Study** as the type of Study you are creating.
5. Enter a name for the Study in the **Name** field.
6. Enter a description for the Study in the **Description** field.
7. Select the **Project** that this Study relates to in the **Project ID** drop down field.
8. Click the **Create Study** button.

Once you have created a Study you can start to upload files to it. For more information on uploading files, click here.

Studies Page

The Studies page provides a central location for creating Studies, uploading files, managing Study permissions, and creating new Workspaces.

Changing organization-categorized Study permissions

To change permissions for an organization's Study:

1. Click the **Permissions** button for the Study.
2. Click the pencil icon beside the Users level.
3. Update permissions as needed.

Note: Changing a user's permissions for external Study results in all existing workspaces for that user losing all access to the Study. This will be addressed in a future release.

Creating a Study and uploading files

To create a Study:

1. Click **Create Study**.
2. Enter a unique ID, type (My study or Organization Study), Name, Description, and Project ID.
3. Click **Create Study**.

To upload files:

1. Click **Upload Files**.
2. You can upload files either by dragging and dropping, or by clicking the **Upload Files** or **Upload Folder** button.

Note: The **Upload Files** button is not visible if you lack write permissions to the Study, or if the Study is external.

Sharing a Study

Organizational Studies can be shared with other members in the Organization. Owners of a Study can amend the permissions of the Study to grant other user's access. There are two permissions available for a Study:

Role	Description
Admin	Able to access the Study, upload files to the Study and grant access to other users.
Readonly	Read only to access to the Study.

To grant a user access to a Study, follow these steps:

1. Ensure that the user you are logged in with has *Admin* permissions on the Study you want to modify. this can be the user who originally created this study or a system administrator.
2. In the portal navigate to the *Studies* page using the menu on the left.
3. Click on the *Organization* or the *My Studies* tab.
4. Find the Study you want to modify the permissions for and expand the *Permissions*section.
5. Next to the *Users* column header, click the edit button.
6. From the drop-down fields that appear, select the required user in either the *Admin* drop down or the *Readonly* drop down field.
7. Click the *Submit* button.

Data Sources: Introduction

Service Workbench can host Studies internally, and provide access to External Studies residing in S3 buckets external to the AWS account where the application was deployed. The Data Sources page enables administrators to configure and manage these Studies.

The external study account must provide Service Workbench with permissions to access the specific S3 bucket and path that contain the data for the External Study. These permissions are created in the external account using an AWS CloudFormation template generated by the application.

The Data Sources page lists external Studies that have been registered with Service Workbench. External Studies also appear on the Studies page, where permissions are configured.

Registering an external Study with Service Workbench

To register an external Study, follow these steps:

Step 1: Set up an external Study account

1. Navigate to the **Data Sources** page using the menu on the left.
2. Click the **Register Studies** button.
3. Enter the AWS Account ID for the AWS account containing the S3 bucket.
    - Previously registered external Study accounts are available in the drop-down.
    - If a previously registered account is chosen, skip to step 2.
4. Select the region that will be used to deploy the onboarding template.
5. Enter an account name for identifying this account in the Service Workbench UI.
6. Specify optional contact information for the account.

Step 2: Specify the S3 bucket details

1. For **Bucket Name**, choose a name of the external S3 bucket.
    - Previously registered external buckets are available in the drop-down.
    - If a previously registered bucket is chosen, skip to step 3.
2. For **Bucket Region**, choose a region.
3. For **Bucket Default Encryption**, if the external bucket uses **AWS Key Management Service key (SSE-KMS)**, then a value is required for the **KMS Arn** field.

Step 3: Specify the Study project details

1. Click **Add Study**.
2. For **Study Id**, enter a unique Study ID.
3. For **Study Name**, enter the Study name.
4. For **Study Folder**, enter the folder name (path in the S3 bucket).
5. For **Project**, choose a project.
6. Choose the **Type**. For more information, refer to Creating a Study.
7. For **Access**, choose either **Read Only** or **Read/Write**. Appropriate permissions can be assigned on a per-user basis on the Studies page, if it is an organization's Study.
8. For **Description**, enter the Study project details.
9. For **Study KMS ARN**, enter the value, if applicable.
10. For **Admin**, choose one or more admins for the Study, if it is an organization's Study. This can be edited later on the **Studies** page.
11. Click **Save & Continue**.

An information panel is displayed indicating the account, bucket, and Study have been registered within Service Workbench. The final step is to use the generated CloudFormation template to onboard the account.

Onboarding an external Study account

If an external study account is being onboarded with Service Workbench for the first time, then *Create Stack* option is selected. The application will generate a new CloudFormation template suitable for first-time onboarding.

If the external study account has already been onboarded (refer to "Registering an External Study with Service Workbench" section above), then the *Update Stack* option is selected. The application generates an update to the previously deployed CloudFormation template.

To onboard an external study account for the first time:

1. In the **Register Studies** window, click **Next** to display CloudFormation template information.
2. The generated CloudFormation template, specific to the external study account and external S3 bucket is displayed. The appropriate *Create Stack* or *Update Stack* option is selected.
3. If you have admin access to the external study account:
    i. In a new browser tab, log in to the external study account in the AWS Management Console. Ensure the correct region is selected.
    ii. In Service Workbench, click the **Create Stack** or **Update Stack** button as appropriate to load the CloudFormation template into the AWS Management Console.
    iii. In AWS Management Console, follow the prompts and click **Create Stack**.
    iv. Click **Done**.
4. If you do not have admin access to the AWS account where the S3 bucket resides:
    i. Click the button to copy the CloudFormation template link to the clipboard.
    ii. Create an email to the admins of the account containing the link to the CloudFormation template. Note: Link to the CloudFormation template is valid for 12 hours.
    iii. Click **Done**.
5. The Data Sources page is displayed, with the newly registered study in the *Pending*status. After all onboarding has been completed and Service Workbench can reach the Study, it will show the status as Available.
6. Click **Test Connection**.

Removing External Studies

Removing external Studies is not supported at this time and it will be available in a future release.

To remove access to the data in the external study account, delete the CloudFormation stack in the AWS account (specified in the "Onboarding an external Study account section above"). This removes permissions enabling access to the data.

Workspaces: Introduction

Service Workbench enables organizations to provide researchers with a centralized location to search for data sets and deploy research workspaces. Researchers can access a portal, quickly find data they are interested in, and quickly begin analysis in SageMaker Notebooks, for example.

Service Workbench also allows an organization to provide access to their data sets, or a subset of their data sets, to external organizations in a controlled way. In addition, the external organization can use their own AWS account for the research workspace and access the data in the hosting organization.

Once a User has found the Study or Studies that they are interested in researching, they can deploy a Workspace with the desired data attached.

A Workspace environment contains a set of tools to access and integrate data. The following environments are currently provided:

- *SageMaker Notebook* – A SageMaker Jupyter Notebook with TensorFlow, Apache MXNet and Scikit-learn2
- *EMR* – An Amazon EMR research workspace with Hail 0.2, JupyterLab, Spark 2.4.4 and Hadoop 2.8.52
  Note: EMR workspaces are not available if AppStream is enabled for the deployment.
- *EC2 - Linux* – An EC2 Linux instance.
- *EC2 - Windows* – An EC2 Windows instance.
- *EC2 - RStudio* – An EC2 RStudio instance.

Note: There's currently a 10K limit on the number of workspaces that can be created in one SWB environment.

Creating a new Workspace

A user can select a Study or multiple Studies and launch a Workspace to access and analyze that data. To launch a Workspace, follow these steps:

1. In the portal navigate to the *Studies* page using the menu on the left.
2. Select the Studies to be attached to the new Workspace.
   i. Note: If the Study is in a Pending or Error state, there may be a permissions issue with the AWS account hosting the data. You will not be able to select the Study. Contact your administrator.
3. Once you have selected all the Studies you want, click the *Next* button.
4. Choose the type of Workspace you want and click the *Next* button.
5. Type a name for the Workspace in the *Name* field.

6. Select a project that this Workspace will belong to in the *Project ID* drop down field.
7. Select the *Configuration* type.
8. Type a description for the Workspace in the *Description* field.
9. Click the *Create Research Workspace* button.

This will deploy the new Workspace and attach the Studies that were selected. You will automatically be redirected to the Workspaces tab on the portal.

NOTE

If you are deploying an EMR, EC2 - Linux or EC2 - Windows based Workspace, you will also be asked to provide a *Whitelisted CIDR*.
Your current IP address is automatically detected as x.x.x.x/32.
This will be used to configure the Security Group associated with this instance enabling you to get access.

TIP

It is recommended that you limit the Whitelisted CIDR range to only the IP addresses that need to access the Workspace.

Accessing a Workspace

You can connect to Workspaces that you have access to. To access a Workspace, follow these steps:
In the portal navigate to the *Workspaces* page using the menu on the left. In the list of Workspaces, find the Workspace that you want to connect to.

Access restrictions: If the restrictAdminWorkspaceConnection flag is enabled, an administrator can only connect to the Workspaces they are assigned to.

Connect to SageMaker and EMR workspace

Click on the *Connect* button, the Workspace must be in the *Ready* state to access it. The selected studies will show up as mounted directories in the Jupyter notebook running on the workspace (EMR or SageMaker). These study directories will contain files uploaded to the corresponding study. Any files uploaded to the study from the Service Workbench will automatically appear in the mounted study directories after a short delay.

Note: The password for EMR instance is 'go-research-on-aws'.

To change the default password for Jupyter Notebook instances, contact your Solution Architect, raise an AWS support case, or follow these instructions:

1. Generate a SHA1 hash for your password choice.

2. Locate line 51 in `main/solution/machine-images/config/infra/provisioners/provision-hail.sh`: `s/sha1:<salt1>:<hash1>/sha1:<salt2>:<hash2>/`
3. Update `<salt2>` and `<hash2>` to match your password's corresponding values.
4. On your local repo, navigate to `main/solution/machine-images`.
5. Run `pnpx sls build-image -s <stage>` to create a new AMI for EMR environment types.
6. Use the generated AMI ID in the environment type configuration key `AmId`. Your selected password becomes active.

Connect to EC2 Linux

1. Click the connections button shown in EC2 Linux instance
2. Click Create Key button, download the key file (This is the only chance to download the key file)
3. Save the key file locally and run **chmod 600** to restrict access to the key file
4. Click 'Use this SSH Key' button and follow the instructions to link to EC2 instance
5. If the 60 seconds count down on the page times out, simply click 'Use this SSH Key' button again and continue
6. SSH to the EC2 Linux machine using the command shown on the screen. Note that you may need to adjust the path of the private key on your local machine.
7. Once you SSH, the selected studies will show up as mounted directories on the EC2 Linux instance. These study directories will contain files uploaded to the corresponding study. Any files uploaded to the study from the Service Workbench will automatically appear in the mounted study directories after a short delay.

Connect to EC2 Windows

Click the connections button, follow the instruction to link to the instance using a local RDP client.
Note: A warning message may pop up for EC2 certificate. This is a normal behavior as the EC2 Windows instance has self-signed SSL cert. Click continue to get connected.
Once you RDP, the selected studies will show up as directories on the EC2 Windows instance in "D" drive. These study directories will contain files uploaded to the corresponding study.
For EC2 Windows, the selected study data is copied to the attached EBS volumes as opposed to being FUSE mounted in case of other workspace types. If the selected study is writeable, the local changes are synchronized back to S3 as soon as possible.
It uses a custom S3 Synchronizer tool (i.e., **c:\workdir\s3-synchronizer.exe**) tool to sync changes from S3 to local EBS volumes and vice versa.
Please be aware of the following limitations specific to EC2 Windows Workspace Types:

Limitations

S3 to Local Sync Limitations

- If the selected study is Read-Only, any changes made under the locally mapped study directory and it's subdirectories will be *LOST* after the periodic sync. No local changes will persist.
- There will be delay of at least the duration equal to the periodic download interval plus the download time for the S3 changes to reflect on local EBS volumes.
- Deleting a subdirectory in studies S3 location will leave the corresponding subdirectory as empty directory on local EBS volume.

Local to S3 Sync Limitations

- Will not upload changes from local to S3 if there is no change in file size (bytes)
- Will not upload changes from local to S3 if the file is empty (i.e., zero bytes)
- *Conflict resolution is undefined:* i.e., if a file is modified in S3 and locally at the same time, the behavior is undefined. Whichever change gets synchronized first may win.

S3 Synchronizer tool

- The synchronizer is automatically started when the EC2 Windows instance is launched
- You can check if S3 Synchronizer tool is running or not by looking for **s3-synchronizer.exe** in Windows Task Manager.
- *Stopping:* To stop the synchronizer, right click on the **s3-synchronizer.exe** in Windows Task Manager and select **End task.**
- *Starting:* To start the synchronizer, run the powershell script **c:\workdir\start-s3-synchronizer.ps1** (right click, select **Run with Powershell**).
- *Troubleshooting:* View log files **c:\workdir\s3-synchronizer-stderr.txt** and **c:\workdir\s3-synchronizer-stdout.txt**

Connect to RStudio

You can connect to RStudio workspace type by using the template and AMI provided in AWS partner's repository. For more information about new RStudio enhancements, refer to the **Create RStudio ALB workspace** section of *Service Workbench Post Deployment Guide*.

Start and Stop workspace

EC2 Windows, EC2 Linux, RStudio and SageMaker workspaces can be stopped when not in use. Click the stop button to stop the workspace, and click the start button to start the workspace again.

Common connection issues

Connection to workspace is restricted to specific CIDR block.

- Check if your public IP is covered by the restricted CIDR block of the workspace.
- Check if workspace type configuration has hard-coded value in field 'AccessFromCIDRBlock'. (Admin only)
- If you're using VPN, your public IP address might change. Try disconnect VPN, and then connect to workspace.

Terminating a Workspace

When you no longer need a workspace, you can terminate it. Follow these steps to terminate a workspace:

1. In the portal navigate to the *Workspaces* page using the menu on the left.
2. In the list of Workspaces, find the Workspace that you want to terminate.
3. Click on the *Terminate* button, the Workspace must be in the *Ready* state to terminate it.

Administrator's view

Dashboard

The dashboard is visible to all users that log in to the solution. It is also the first screen visible to a user after logging into the web interface. Based on the role a user holds they can see more or less data inside the dashboard's widgets.
The dashboard displays:

- The *Index* costs for the past 30 days. This is the cost; a certain index or cost center has allotted over the course of the last rolling 30 days. This includes all workspaces started for this index, regardless of the user who started or owns them.
- The *Workspaces* costs for the past 30 days. This represents the workspaces owned by the specific, logged in, user over the course of the last 30 days.
- Yesterday's *Workspaces* costs.
- A breakdown of the *Index* costs for the past 30 days per user.

Auth Introduction

The solution provides the ability to connect to Active Directory as an Identity Provider. The web-based user interface does not provide the ability to add or update these settings. These settings are defined as part of the deployment of the solution.
Refer to the ***Service Workbench Configuration Guide*** for more information.

Users Introduction

You can manage Users and Roles in the solution from the *Users* page. Users are entities that can access the solution and Roles define the permissions for Users.

User Roles

There are 5 different types of Roles that you can assign to Users:

Role	**Description**
Admin	Full access to the solution
Internal Guest	Read Only access to the solution
External Guest	Read Only access to the solution
Internal Researcher	Search for Studies, Upload Personal Studies and Start Workspaces
External Researcher	Search for Studies, Upload Personal Studies and Start Workspaces in an External Account

Adding Federated Users

Add a single federated user

To add a single federated user, follow so that they can access the solution, follow these steps:

1. In the portal navigate to the *Users* page using the menu on the left.
2. Click on the *Add Federated User* button.
3. In the *Username* field, enter the email address of the user you are adding.
4. Select the identity provider the user belongs to in the *Identity Provider* field.
5. Select the desired *User Role* from the *User Role* field.
6. Select if the user should be Active or Inactive for the *Status* field.

Add multiple federated users

You can upload a CSV file to create multiple users. The CSV file requires the following columns specified:

- email
- userRole
- identityProviderName

Once you have your CSV file, follow these steps:

1. In the portal navigate to the *Users* page using the menu on the left.
2. Click on the *Add Federated User* button.
3. Click on the *Add Multiple Users* tab along the top.
4. Either Drag and Drop the CSV file of users into the Drag and Drop area, or click on the Drag and Drop area to select the CSV file you want to upload.
5. Click the *Submit* button.

Accounts

Projects Introduction

A Project is a logical grouping of resources that users have access to. In order for a Researcher to upload their own Study data sets and to deploy and user Workspaces a user must have access to a Project.

Projects must be associated to an **Index**. A Project can only be associated to single **Index**.

Create a New Project

To create a new project, follow these steps:

1. In the portal navigate to the *Accounts* page using the menu on the left.

2. Click on *Projects* tab along te top.
3. Click on the *Add Project* button.
4. Type an ID for the project in the *Project ID* field, this is a free text field.
5. Select the *Index ID* that this Project is to be associated with.
6. Optionally provide a *Description*.
7. Click the *Add Project* button.

Adding a User to a Project

You can add *Users* to a Project. This enables that user to access *Studies* and deploy *Workspaces* in the Project. To add a User to a Project, follow these steps:

1. In the portal navigate to the *Users* page using the menu on the left.
2. Click the *Detail* button for the User you want to add the Project to.
3. Click the *Edit* button.
4. From the *Project* field, select the projects that you want to grant access to.
5. Click the *Save* button.

Indexes Introduction

Indexes are a mechanism of grouping **Projects** together for cost reporting purposes. There can be many **Projects** associated with a single Index. There can only be a single AWS Account associated with an Index.

Create a New Index

You can create a new Index that can then be used for multiple projects. To create a new Index, follow these steps:

1. In the portal navigate to the *Accounts* page using the menu on the left.
2. Click on *Indexes* tab along the top.
3. Click on the *Add Index* button.
4. Type an ID for the Index in the *Index ID* field, this is a free text field.
5. Select the *AWS Account ID* that this Index is to be associated with.
6. Optionally provide a *Description*.
7. Click the *Add Index* button.

AWS Accounts

Introduction to AWS Accounts

There are 3 types of AWS accounts roles in this solution.

- The *Main* AWS account. This is the account that this solution is running in.

- The *Master* AWS account. This (optional) account can be the same as the Main AWS account or a different one. It hosts AWS Organizations, needed if the AWS account creation functionality is required.
- The *Member* (or researcher) account(s). These accounts are where researcher analytics are run. These accounts are either member accounts in the organization ('created' accounts) or standalone accounts ('invited' accounts).

Create an AWS Account

When you attempt to create a *Member AWS Account*, the *Main AWS Account* will assume a role in the *Master AWS Account*. Once the role has been assumed, it will then create a *Member AWS Account*.
Once the *Member AWS Account*has been created, the *Main AWS Account* will assume a role in that account and launch a CloudFormation template to build resources (VPC, Subnet, *Cross Account Execution Role*).
To create a new *Member AWS Account,* follow these steps:

1. In the portal navigate to the *Accounts* page using the menu on the left.
2. Click the *AWS Accounts* tab along the top.
3. Click the *Create AWS Account* button.
4. Type a name for the AWS Account in the *Account Name* field.
5. Type an email address for the AWS Account in the *AWS Account Email* field.
6. Provide the *Master Role* ARN for the *Master AWS Account* in the *Master Role Arn* field.
7. Type the External ID for the AWS Account in the *External ID* field.
8. Type a description for the AWS Account in the *Description* field.

Invite an AWS Member Account

As well as *Creating a Member Account*, you can also invite an AWS Account to the solution. The AWS Account that is being invited will need to provide a VPC, one Subnet from within it, and a *Cross Account Execution Role*.

Adding an AWS Account

To add an AWS Account to the solution, follow these steps:

1. In the portal navigate to the *Accounts* page using the menu on the left.
2. Click on the *AWS Accounts* tab along the top.
3. Click the *Add AWS Account* button.
4. Type a name for the AWS Account in the *Account Name* field.
5. Provide the AWS Account ID (12 digits) in the *AWS Account ID* field.

6. Provide the Role Arn that has created when creating a *Cross Account Execution Role* in the *Role Arn* field.
7. Enter the External ID for the AWS Account in the *External ID* field.
8. Provide the VPC ID that *Workspaces* will be deployed into in the *VPC ID* field.
9. Provide the KMS Encryption Key ARN for the AWS Account in the *KMS Encryption Key ARN* field.

Set AWS Account Budget

To set budget for an AWS account, follow these steps:

1. In the portal navigate to the *Accounts* page using the menu on the left.
2. Click on the *AWS Accounts* tab along the top.
3. Click the *Budget Detail* button.
4. Provide Budget Limit, start date and end date (End date needs to be less than a year from start date)
5. Provide notification thresholds and email
6. Thresholds and email are interdependent, please fill them both to get notification or remove both to turn off notification

Once the budget is set, AWS Budget will be monitoring the actual spent of the corresponding AWS account. Alert will be sent to the notification email address when the actual spent breach each of the thresholds.
Alert will also be sent 7 days prior to budget end date. If notification email is set, budget end date alert will be sent to the notification email. if notification email is not set, budget end date alert will be sent to the email address the AWS account is registered under.

Roles and Permissions

Master Account Role

This role resides in the master AWS account and is assumed by the main AWS account.

Master Role Trust Policy

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::${MainAccount}:root"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "sts:ExternalId": ${ExternalId}
        }
      }
    }
```

```
  ]
}
```

Master Role Permissions

The following details the Managed and Inline Policy permissions needed.

Managed Policy: AWSOrganizationsFullAccess

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "organizations:*",
            "Resource": "*"
        }
    ]
}
```

WARNING: You should restrict the actions to **createAccount**, **describeCreateAccountStatus**and **describeAccount** only.

Inline Policy: sts:AssumeRole

This policy is for the controlling role between master AWS account and master AWS account:

```
{
    "Version": "2012-10-17",
    "Statement": {
        "Effect": "Allow",
        "Action": "sts:AssumeRole",
        "Resource": "arn:aws:iam::*:role/OrganizationAccountAccessRole"
    }
}
```

Cross Account Execution Role

This role resides in the *Member AWS Account(s)* and is assumed by the *Main AWS Account*. When *Creating a Member AWS Account(s)* in the organization of the *Master AWS Account(s)*, this role is created by the **solution/packages/cfn-templates/lib/templates/onboard-account.yaml** template.

Trust Policy

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "arn:aws:iam::${MainAccount}:role/${ApiHandlerRole}",
          "arn:aws:iam::${MainAccount}:role/${WorkflowLoopRunnerRole}",
          "arn:aws:iam::${MemberAccount}:root"
        ]
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "sts:ExternalId": ${ExternalId}
```

```
        }
      }
    }
  ]
}
```

The principals listed above are:

- ApiHandlerRole: A role in the Main AWS account associated with the Service Workbench backend API execution.
- WorkflowLoopRunnerRole: A role in the Main AWS account associated with background workflow execution as initiated by backend API calls.
- The Member AWS account itself.

Cross Account Execution role permissions

These policies support running analytics.

CloudFormation

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "cloudformation:CreateStack",
        "cloudformation:DeleteStack",
        "cloudformation:DescribeStacks",
        "cloudformation:DescribeStackEvents"
      ],
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

Cost Explorer

```
{
  "Statement": {
    "Action": ["ce:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
}
```

TIP

You will need to ensure that *Cost Explorer* has been enabled in the member account. See here for more information.

EC2

```
{
  "Statement": {
    "Action": ["ec2:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
```

```
}
```

EMR

```
{
  "Statement": {
    "Action": ["elasticmapreduce:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
}
```

IAM

```
{
  "Statement": {
    "Action": ["iam:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
}
```

S3

```
{
  "Statement": {
    "Action": ["s3:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
}
```

SageMaker

```
{
  "Statement": {
    "Action": ["sagemaker:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
}
```

SSM

```
{
  "Statement": {
    "Action": ["ssm:*"],
    "Resource": "*",
    "Effect": "Allow"
  }
}
```

Devops Account Role & Trust Policy

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": [
```

```
                    "arn:aws:iam::${MainAccount}:role/${ApiHandlerRole
}",
                    "arn:aws:iam::${MainAccount}:role/${WorkflowLoopRu
nnerRole}"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "sts:ExternalId": ${ExternalId}
                }
            }
        }
    ]
}
```

Devops Role Permissions

AllowAMISharingtoSWBMainAccount

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "ec2:ModifyImageAttribute",
                "ec2:DescribeImages",
                "appstream:UpdateImagePermissions"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "EC2AMISharingAccess"
        }
    ]
}
```

Studies: Introduction

Service Workbench provides multiple ways to connect compute workspaces to data (Studies) saved in Amazon S3. As a researcher, you can also use it to search data using tools specific to your requirements.

Using Service Workbench, an organization can share Studies with other organizations with access controls.

There are three types of Studies available in Service Workbench. You can mount multiple Studies, of any type, to your Workspaces.

Type	Description
My Studies	Specifies Studies created by users. Use this option to work on datasets that are exclusive to you or that are used specifically for your research.
Organization Studies	Specifies Studies that can be shared with other users. It contains data that had been collected by an organization or is licensed. You can grant or deny users access to this data in order to comply with regulations or licensing restrictions on the data.
Open Data	Service Workbench provides access to Open Data on AWS data by frequently scanning the set of AWS open datasets and adding new datasets to this category. This can include 1,000 genomes and other datasets openly available through Amazon.

Service Workbench can host data for My Studies and Organization Studies internally, in the studydata S3 bucket created in the AWS account where Service Workbench was deployed. The application can also provide access to Studies hosted in external S3 buckets in other AWS accounts using the **Datasets** page.

Creating a Study

Studies are data sets. Admin, Internal Researchers and External Researchers can search for Studies to perform research on. They can also create new Studies and upload data to Studies. To create a new Study, follow these steps:

1. In the portal navigate to the *Studies* page using the menu on the left.
2. Click the *Create Study* button.
3. Provide an ID for the Study in the *ID* field.
4. Choose either *My Study* or *Organizational Study* as the type of Study you are creating.
5. Enter a name for the Study in the *Name* field.
6. Enter a description for the Study in the *Description* field.
7. Select the *Project* that this Study relates to in the *Project ID* drop down field.
8. Click the *Create Study* button.

Once you have created a Study you can start to upload files to it. For more information on uploading files, click here.

Studies Page

The Studies page provides a central location for creating Studies, uploading files, managing Study permissions, and creating new Workspaces.

Changing organization-categorized Study permissions

To change permissions for an organization's Study:

1. Click the *Permissions* button for the Study.
2. Click the pencil icon beside the Users level.
3. Update permissions as needed.

Note: Changing a user's permissions for an external Study result in all existing workspaces for that user losing all access to the Study. This will be addressed in a future release.

Creating a Study and uploading files

To create a Study:

1. Click *Create Study*.
2. Enter a unique ID, type (My study or Organization Study), Name, Description, and Project ID.
3. Click *Create Study*.

To upload files:

1. Click *Upload Files*.
2. You can upload files either by dragging and dropping, or by clicking the *Upload Files* or *Upload Folder* button.

Note: The *Upload Files* button is not visible if you lack write permissions to the Study, or if the Study is external.

Sharing a Study

Organizational Studies can be shared with other members in the Organization. Owners of a Study can amend the permissions of the Study to grant other user's access. There are two permissions available for a Study:

Role	**Description**
Admin	Able to access the Study, upload files to the Study and grant access to other users.
Readonly	Read only to access to the Study.

To grant a user access to a Study, follow these steps:

1. Ensure that the user you are logged in with has *Admin* permissions on the Study you want to modify. this can be the user who originally created this study or a system adminsitrator.
2. In the portal navigate to the *Studies* page using the menu on the left.
3. Click on the *Organization* or the *My Studies* tab.
4. Find the Study you want to modify the permissions for and expand the *Permissions*section.
5. Next to the *Users* column header, click the edit button.

6. From the drop-down fields that appear, select the required user in either the *Admin* drop down or the *Readonly* drop-down field.
7. Click the *Submit* button.

Data Sources

Introduction

Service Workbench can host Studies internally, and provide access to External Studies residing in S3 buckets external to the AWS account where the application was deployed. The Data Sources page enables administrators to configure and manage these Studies.

The external study account must provide Service Workbench with permissions to access the specific S3 bucket and path that contain the data for the External Study. These permissions are created in the external account using an AWS CloudFormation template generated by the application.

The Data Sources page lists external Studies that have been registered with Service Workbench. External Studies also appear on the Studies page, where permissions are configured.

Registering an external Study with Service Workbench

To register an external Study, follow these steps:

Step 1: Set up an external Study account

1. Navigate to the *Data Sources* page using the menu on the left.
2. Click the *Register Studies* button.
3. Enter the AWS Account ID for the AWS account containing the S3 bucket.
    - Previously registered external Study accounts are available in the drop-down.
    - If a previously registered account is chosen, skip to step 2.
4. Select the region that will be used to deploy the onboarding template.
5. Enter an account name for identifying this account in the Service Workbench UI.
6. Specify optional contact information for the account.

Step 2: Specify the S3 bucket details

1. For *Bucket Name*, choose a name of the external S3 bucket.
    - Previously registered external buckets are available in the drop-down.
    - If a previously registered bucket is chosen, skip to step 3.
2. For *Bucket Region*, choose a region.
3. For *Bucket Default Encryption*, if the external bucket uses *AWS Key Management Service key (SSE-KMS)*, then a value is required for the *KMS Arn* field.

Step 3: Specify the Study project details

1. Click *Add Study*.
2. For *Study Id*, enter a unique Study ID.
3. For *Study Name*, enter the Study name.
4. For *Study Folder*, enter the folder name (path in the S3 bucket).
5. For *Project*, choose a project.
6. Choose the *Type*. For more information, refer to Creating a Study.
7. For *Access*, choose either *Read Only* or *Read/Write*. Appropriate permissions can be assigned on a per-user basis on the Studies page, if it is an organization's Study.
8. For *Description*, enter the Study project details.
9. For *Study KMS ARN*, enter the value, if applicable.
10. For *Admin*, choose one or more admins for the Study, if it is an organization's Study. This can be edited later on the *Studies* page.
11. Click *Save & Continue*.

An information panel is displayed indicating the account, bucket, and Study have been registered within Service Workbench. The final step is to use the generated CloudFormation template to onboard the account.

Onboarding an external Study account

If an external study account is being onboarded with Service Workbench for the first time, then Create Stack option is selected. The application will generate a new CloudFormation template suitable for first-time onboarding.

If the external study account has already been onboarded (refer to "Registering an External Study with Service Workbench" section above), then the Update Stack option is selected. The application generates an update to the previously deployed CloudFormation template.
To onboard an external study account for the first time:

1. In the *Register Studies* window, click *Next* to display CloudFormation template information.
2. The generated CloudFormation template, specific to the external study account and external S3 bucket is displayed. The appropriate Create Stack or Update Stack option is selected.
3. If you have admin access to the external study account:
    i. In a new browser tab, log in to the external study account in the AWS Management Console. Ensure the correct region is selected.
    ii. In Service Workbench, click the *Create Stack* or *Update Stack* button as appropriate to load the CloudFormation template into the AWS Management Console.
    iii. In AWS Management Console, follow the prompts and click *Create Stack*.

iv. Click *Done*.

4. If you do not have admin access to the AWS account where the S3 bucket resides:
   i. Click the button to copy the CloudFormation template link to the clipboard.
   ii. Create an email to the admins of the account containing the link to the CloudFormation template. Note: Link to the CloudFormation template is valid for 12 hours.
   iii. Click *Done*.
5. The Data Sources page is displayed, with the newly registered study in the Pendingstatus. After all onboarding has been completed and Service Workbench can reach the Study, it will show the status as Available.
6. Click *Test Connection*.

Removing External Studies

Removing external Studies is not supported at this time and it will be available in a future release.

To remove access to the data in the external study account, delete the CloudFormation stack in the AWS account (specified in the "Onboarding an external Study account section above"). This removes permissions enabling access to the data.

Best Practices

Multiple Deployment Environments

It is recommended that you deploy separate 'Dev', 'Test' and 'Production' environments of this solution.

It is also possible to deploy each environment to different AWS accounts for further separation.

Amazon Inspector

Amazon Inspector tests the network accessibility of your Amazon EC2 instances and the security state of your applications that run on those instances. Amazon Inspector assesses applications for exposure, vulnerabilities, and deviations from best practices. After performing an assessment, Amazon Inspector produces a detailed list of security findings that is organized by level of severity.

You can manually configure Amazon Inspector to monitor the Research Workspaces that you deploy. For more information see the *Amazon Inspector User Guide*.

AWS CloudTrail

AWS CloudTrail is an AWS service that helps you enable governance, compliance, and operational and risk auditing of your AWS account. Actions taken by a user, role, or an AWS service are recorded as events in CloudTrail. Events include actions taken in the AWS Management Console, AWS Command Line Interface, and AWS SDKs and APIs.

CloudTrail is enabled on your AWS account when you create it. When activity occurs in your AWS account, that activity is recorded in a CloudTrail event. You can easily view recent events in the CloudTrail console by going to Event history. For an ongoing record of activity and events in your AWS account, *create a trail*.

AWS Shield

A distributed denial of service (DDoS) attack is an attack in which multiple compromised systems attempt to flood a target, such as a network or web application, with traffic. A DDoS attack can prevent legitimate users from accessing a service and can cause the system to crash due to the overwhelming traffic volume.

AWS provides two levels of protection against DDoS attacks: AWS Shield Standard and AWS Shield Advanced. For production workloads we recommend that you enable and configure AWS Shield Advanced.

For more information see the How AWS Shield works documentation.

CI/CD

We recommend that you setup a CI/CD pipeline for this solution. This section will provide you with an overview of how to do this.

Terminology

- *Source Account:* AWS Account containing CodeCommit repository with the source code
- *Target Account:* AWS Account where the solution needs to be deployed to. The CodePipeline is also deployed in the target account.
- *Staging Environment:* Solution environment created to run integration tests or manual tests before deploying the solution to final target environment.
- *Target Environment:* Target solution environment where the code needs to be deployed.

How the CI/CD pipeline works

At high level the pipeline works as follows:

1. Every commit in the source account on the configured repo for the specified branch triggers a CloudWatch event.
2. A CloudWatch Rule pushes the event to the default event bus of the target account.
3. A CloudWatch Rule in the target account triggers the CodePipeline.
4. The pipeline contains the following stages and executes them in order. The pipeline stops upon failure of any stage and notifies user via configured SNS topic.
    i. *Source:* This stage takes the code from the CodeCommit repository from the specified branch and uploads it to an S3 bucket. This S3 bucket is called artifacts bucket and is used by the CodePipeline to pass artifacts from one stage to other stage.
    ii. *Build-And-Deploy-To-Staging-Env:* This stage uses AWS CodeBuild to perform build and deployment. It downloads the code from the artifacts S3 bucket and installs dependencies, performs the static code analysis, runs unit tests, and deploys to the staging environment. This stage is only created if `createStagingEnv` setting is set to true in settings file. Developers can set `createStagingEnv` to false to skip creation and deployment to staging environment and directly push changes to their target development environment. This flag should be set to true for higher environments (such as **demo** or **production**) to make sure code is deployed and tested in a staging environment before pushing to target environment.
    iii. *Test-Staging-Env:* This stage uses AWS CodeBuild to execute integration tests against the staging environment. This stage is only created if `createStagingEnv` setting is set to `true` in settings file. Developers can set `createStagingEnv` to false to skip creation and deployment to staging environment and directly push changes to their target development environment.
    iv. *Push-To-Target-Env:* This stage is for manual approval to deploy to target environment. The pipeline will pause at this stage and wait for manual approval. The user will receive an email notification via configured SNS topic. The notification email address is configured via the setting `emailForNotifications` in the settings file. This stage is only created if `requireManualApproval` setting is set to true in settings file. Setting `requireManualApproval` to false will cause auto-propagation to the target environment.
    v. *Build-and-Deploy-to-Target-Env:* This stage uses AWS CodeBuild to perform build and deployment. It downloads the code from the artifacts S3 bucket and installs dependencies, performs the static code analysis, runs unit tests, and deploys to the target environment.
    vi. *Test-Target-Env:* This stage uses AWS CodeBuild to execute integration tests against the target environment. This stage is only created if `runTestsAgainstTargetEnv` setting is set to **true** in settings file. Developers can set `createStagingEnv` to `false`, `requireManualApproval` to `false`

, and `runTestsAgainstTargetEnv` to **true** to skip creation and deployment to staging environment and directly push changes to their target development environment without requiring manual approval and run integration tests directly against their target development environment.

Deploying the CI/CD Pipeline

Follow these steps to deploy the CI/CD Pipeline:

1. Deploy the `cicd-source` stack to the Source Account

Start by creating a settings file in `cicd/cicd-source/config/settings` for the environment for which you want to create the CI/CD pipeline. For example, to create the CI/CD pipeline for **dev** environment, create a `dev.yml` file in `cicd/cicd-source/config/settings/`.

You can create the settings file by copying the sample `demo.yml` file. Please adjust the settings as per your environment. Read the inline comments in the file for information about each setting. In the setting file you have just created, set the following settings as **"*"**

```
artifactsS3BucketArn: "*"artifactsKmsKeyArn: "*"
```

You can now deploy the `cicd-source` stack.

```
cd cicd/cicd-sourcepnpx sls deploy --stage <env-name>
```

2. Deploy `cicd-pipeline` stack to the target account.

Create a settings file in `cicd/cicd-pipeline/config/settings` for the environment for which you want to create the CI/CD pipeline. For example, to create the CI/CD pipeline for **dev** environment, create `dev.yml` file in `cicd/cicd-pipeline/config/settings/`. You can create the settings file by copying the sample `demo.yml` file.

Please adjust the settings as per your environment. Read inline comments in the file for information about each setting.

You can now deploy the `cicd-pipeline` stack:

```
cd cicd/cicd-pipelinepnpx sls deploy --stage <env-name>
```

3. Re-deploy cicd-source stack to the source account. You will need to re-deploy the cicd-source stack to the Source Account to lock down permissions for the artifacts bucket.

Note down the CloudFormation stack output variables `AppArtifactBucketArn` and `ArtifactBucketKeyArn` from the `cicd-pipeline` stack you deployed above.

You can also use `sls info` command with `--verbose` flag to print stack output variables from the cicd/cicd-pipeline folder:

```
pnpx sls info --verbose -s <env-name>
```

4. Update CloudFormation variables

Next you will need to set the CloudFormation stack output variables for `AppArtifactBucketArn` and `ArtifactBucketKeyArn` that you obtained in the previous steps for the settings `artifactsS3BucketArn` and `artifactsKmsKeyArn` in settings file `cicd/cicd-source/config/settings/demo.yml`.

```
artifactsKmsKeyArn: "<value of the CloudFormation output variable
ArtifactBucketKeyArn from cicd-pipeline stack>"
```

Next, you will need to re-deploy the `cicd-source` stack to lock down the permissions in `CodeCommitSourceRole`.

```
cd cicd/cicd-sourcepnpx sls deploy --stage <env-name>
```