

ASAP CRN Cloud User Manual



Contents

| | |
|---|----|
| Background | 1 |
| Purpose..... | 1 |
| Getting Started | 1 |
| Evaluating Collections | 1 |
| Searching for Data..... | 2 |
| Harmonized Collections..... | 3 |
| Member Collections | 4 |
| All Collections | 5 |
| Submitting for Data Access | 5 |
| Request Access Form | 5 |
| Guidance for Providing Intended Use | 8 |
| Application Process | 8 |
| Reviewing Available Data Through CRN Cloud as an Approved User..... | 9 |
| Login..... | 9 |
| Review Tables | 10 |
| View Query | 11 |
| Download Data | 12 |
| Analyzing Data | 13 |
| Accessing Verily Workbench | 13 |
| Reference Workspace..... | 13 |
| Verily Workbench Setup | 13 |
| Establishing a Personal Workspace..... | 13 |
| Establishing a Spend Profile | 14 |
| Establishing a Development Environment..... | 15 |
| Accessing Data..... | 16 |
| Navigating Data on Verily Workbench..... | 16 |
| Loading Data Using Saved Searches | 17 |
| Analytical Functionality | 20 |
| Appendix | 22 |
| Data Dictionary | 22 |
| Help and Troubleshooting..... | 22 |
| Resources for Contributing Teams | 23 |
| Authorship List..... | 23 |

Background

Aligning Science Across Parkinson's (ASAP) Collaborative Research Network (CRN) cloud platform, ASAP CRN Cloud, is a platform designed to provide an environment for pre-publication collaboration and responsible data sharing within the CRN. The goal of the platform is to accelerate the pace of discovery for Parkinson's Disease research, which is expected to be accomplished through delivery of the following features:

- Review Descriptions of 'Controlled Access' Data without the need to submit a Data Use Application

Allowing the opportunity to review the descriptions of available data prior to agreeing to the terms of a data use application is expected to allow for increased visibility into available data.

- Harmonized Data Collections

Within harmonized data collections, the data is harmonized across points of commonality.

- Easy Access and Analysis through Verily Workbench

The opportunity to utilize Verily Workbench for conducting analysis of data collections to allow for access to large-volume data and facilitate collaboration in a shared environment.

Purpose

The purpose of the ASAP CRN Cloud User Manual is to provide guidance for how to navigate the tool and make use of the data through Verily Workbench.

Getting Started

Access ASAP CRN Cloud by navigating to cloud.parkinsonsroadmap.org. There is no need to log in to the tool to review descriptions of the available data and harmonized collections. From the landing page you can follow the directions provided in [Evaluating Collections](#) to search the data for related terms or navigate the collections.

Evaluating Collections

The instructions in this section are intended to provide guidance for exploring the data available in ASAP CRN Cloud.

Searching for Data

Locate collections of interest to your research team by using the search function in the toolbar on the top of the page.

To perform a search, enter the term of interest into the field then select enter on your keyboard.

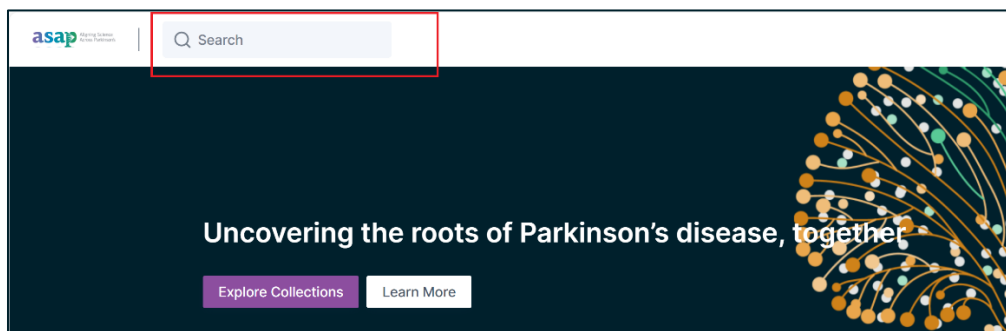


Figure 1 - Searching for Data

The software will run a query using your search term and open the search page, populated with your results.

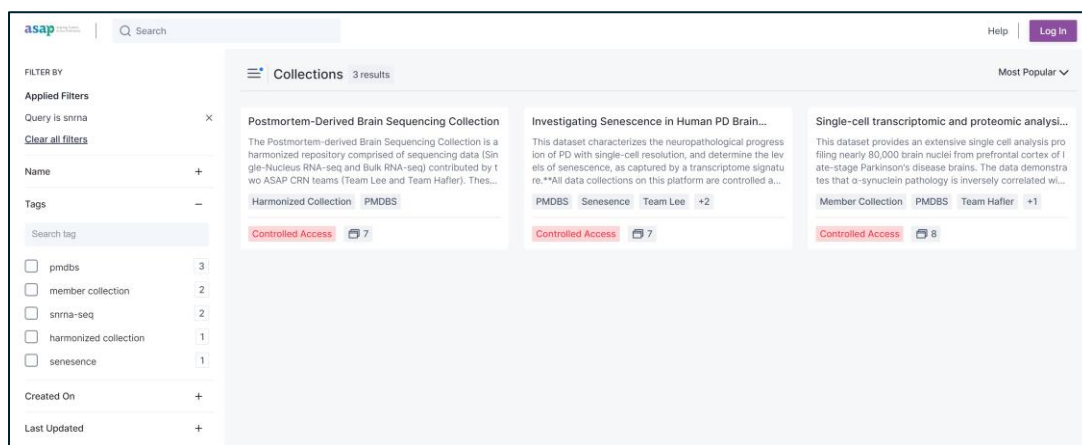


Figure 2 - Search Page Loaded from Search Field

Next, you can open a collection tile by clicking any of the displayed collections or use the menu on the left to filter the results further.

The search field will remain available as you navigate through ASAP CRN Cloud. A search of the available collections can be performed from any page.

Harmonized Collections

Harmonized collections provide the data harmonized across multiple points of commonality to create a more robust dataset. You can review the available harmonized data collections by selecting the 'View Harmonized Collection' link. When you select the link, the tool will open the Post-mortem Derived Brain Sequencing Collection.

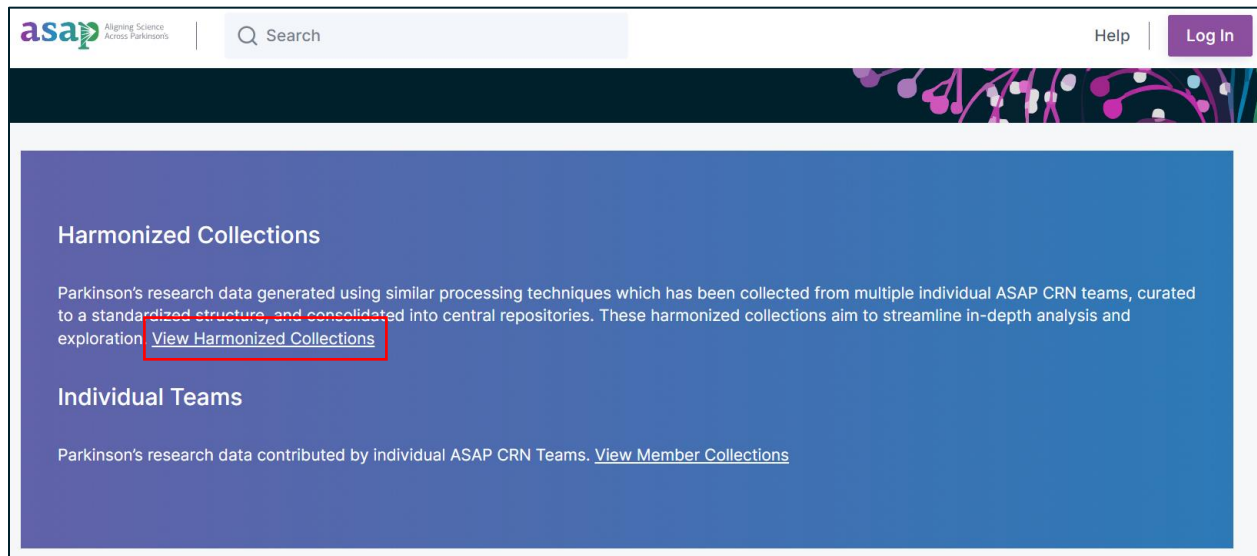


Figure 3 - Select Harmonized Collection from Landing Page

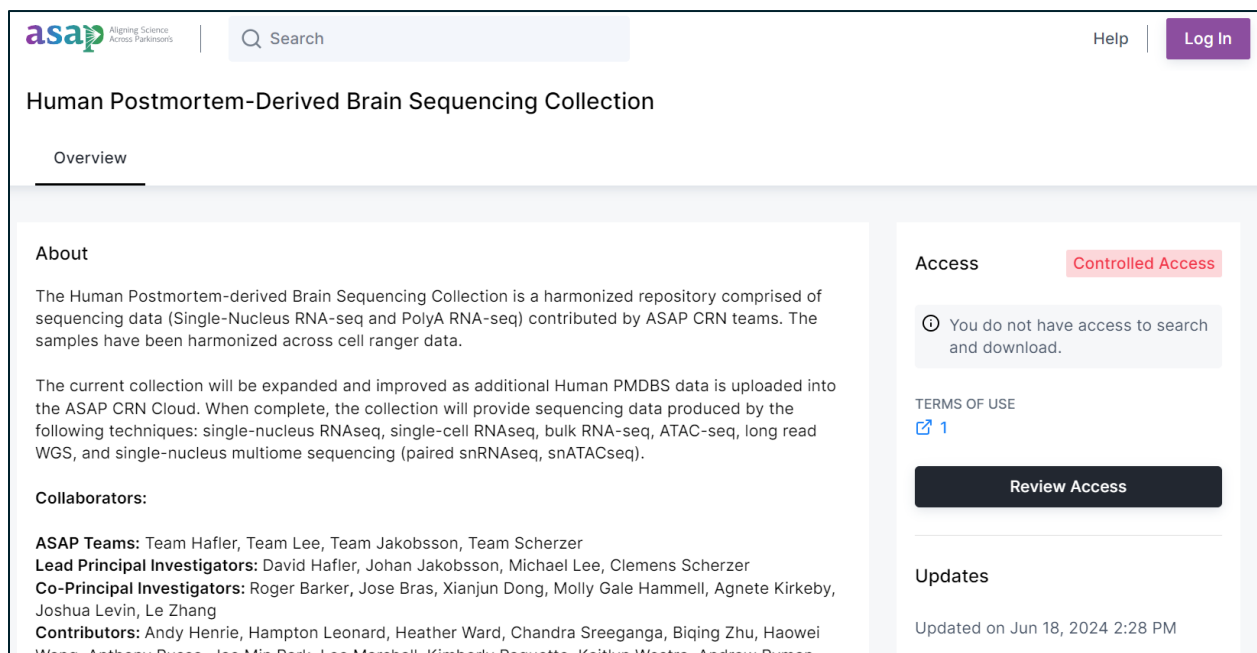


Figure 4 - Post-Mortem Derived Brain Sequencing



Note: Data files will not be displayed until your Data Use Application been approved and you have been granted permissions.

Member Collections

Member collections provide data and background for individual CRN teams. You can review member collections by selecting the 'View Member Collections' link. When you select the link, the tool will open the Search page displaying all member collections.

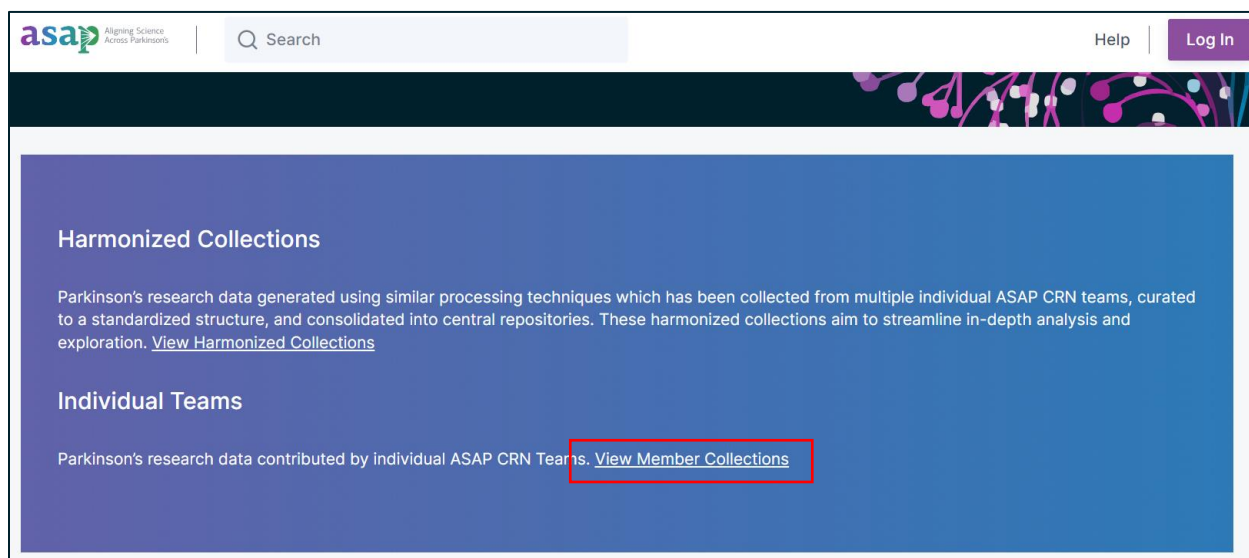


Figure 5 - Select Member Collections from Landing Page

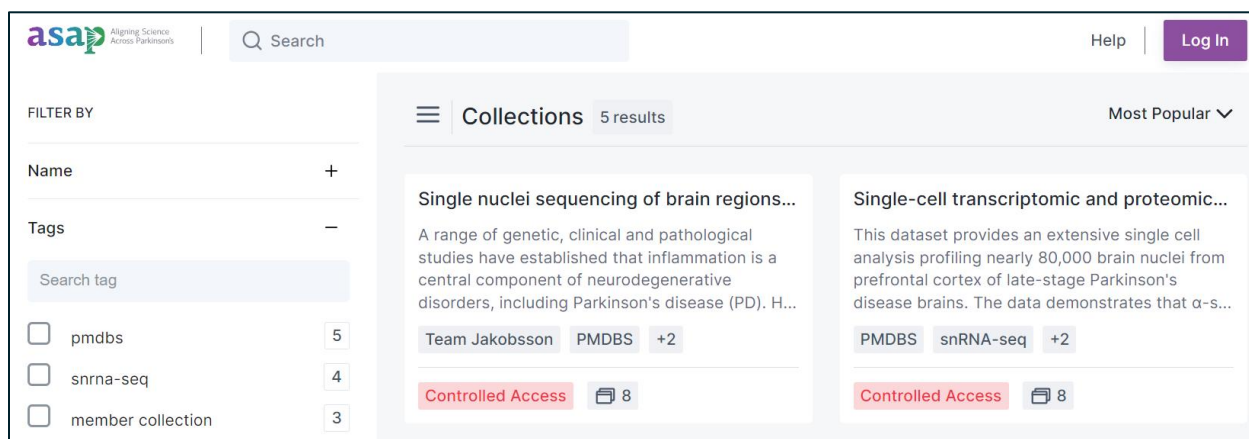


Figure 6 - Member Collections on Search

All Collections

To view a combination of all member and harmonized collections, select the **Explore Collections** button from the ASAP CRN Cloud landing page.

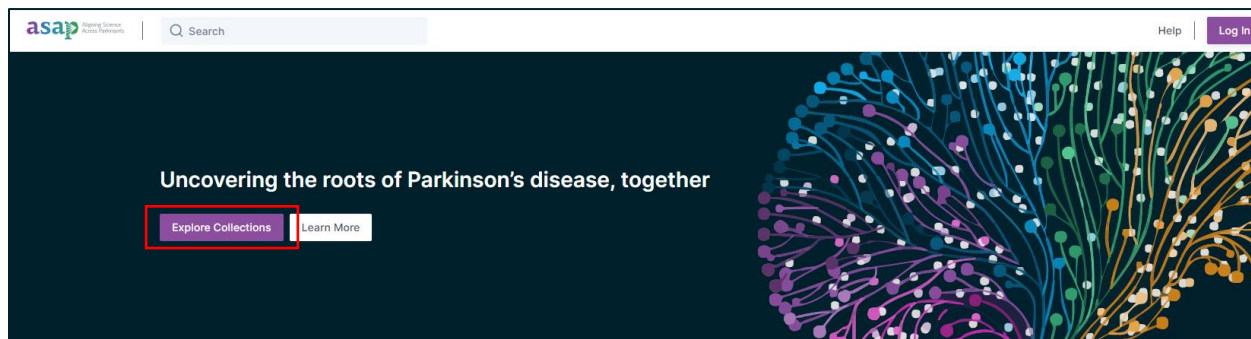



Figure 7 - Select Explore Collections to View All Collections

Submitting for Data Access

Access to data provided through ASAP CRN Cloud requires acceptance of the ASAP CRN Cloud Data Use Agreement and an approved application. The application will typically take 5 business days to review and approve. Once data access has been granted users will have access to all ASAP CRN Cloud data.

Request Access Form



Note: You will need a Google-enabled email address to apply for data access. To enable the email linked to your organization, follow these instructions: [Instructions for Enabling Your Institutional Email](#).

1. Log in to CRN Cloud application by selecting the **Log In** button in the top right-hand corner of the page.
2. To submit for data access, you must first open a collection tile (see [Navigating Collections](#)).
3. From within the collection tile, locate the Access section on the right side of the page. Within the Access section, select the **Review Access** button.

Access

Controlled Access

You do not have access to search and download.

TERMS OF USE

[Version 2.0](#)

Review Access

Figure 8 - Select Review Access to Apply for Data Access

Review Access

Submit a request to access services in this collection.

| Services | Status | Action |
|--|----------------------|-----------------------------|
| Search and Download contents of the collection | <div>No Access</div> | <div>Apply for Access</div> |

Cancel

Figure 9 – Review Access – Select Apply for Access

Apply for Access

Submit a request to search and Download contents of the collection.

Full Name

K

Kate Trimble

Email

knttesting12@gmail.com

Highest Level of Education *

Please enter a value.

Role *

Other Role

If you answered "Other" above, please describe your current role.

Affiliation / Company *

Country *

Cancel

Submit

Figure 10 – Apply for Access Form

- Complete the form by providing the requested information (e.g. – name, institution, intended use). See [Guidance for Providing Intended Use](#) for advice on how to provide an intended use that is likely to be approved.
- When the form is complete, select the **Submit** button to submit application for review.



Note: Please provide sufficient information to evaluate your intended use. See [Guidance for Providing Intended Use](#) for advice on how to provide an intended use that is likely to be approved.

Guidance for Providing Intended Use

The intended use submission should provide context to how you plan to use the data by giving background information about your planned analysis.

Some examples of an excellent explanation of intended use are as follows:

Example 1:

Our group has been generating data to study transcriptional changes in familial risk genes associated with Parkinson's disease. In our analysis, we have compared specific RNA transcript isoforms of these genes and observed differences based on genetic carrier status. We are seeking access to this data to apply an isoform variant pipeline and eQTL approach to the brain single cell data from CRN studies to evaluate whether these transcripts are present, and whether this data type is sufficient to expand our study. We will utilize the data sets from Team X, which were generated using the single cell RNAseq, and will utilize the available WGS data from GP2 to establish genetic carrier status. We will also utilize the transcriptomic data from CRN Team Z to evaluate the presence of these transcripts and see whether allele specific expression is correlated with isoform bias within PD patient brains. We have reached out to these teams regarding the data structure and their ongoing work, and have reviewed the data and methods.

Example 2:

We are performing a mouse screening study to evaluate lysosomal pathways in PD. We are seeking access to the data within the CRN cloud to test whether our result in mouse is supported by human postmortem data. This is an exploratory effort and we believe that two specific genes may serve as critical regulators. A recent paper from ExceptionalPostdoc et al. (2024) reported that up to 5% of the population harbor a variant of unknown significance, and we hope to evaluate whether such carriers are present in these data sets in a pilot effort towards further mechanistic study. We will leverage the genetic data to identify these carriers, and will specifically test a small subset of the participants included from Brain Bank X, based on available metadata.

Examples of what might not be acceptable:

- We are requesting access to the data in the CRN cloud to look at PD brain data to study gene expression in PD.
- I am building a database and wish to access data from ASAP.
- I work with a CRO and wish to analyze this data.
- I am requesting access for my team to conduct research.

Application Process

The application process is expected to take up to 5 business days, but it may take longer if additional review is required. After your application has been reviewed and an approval decision has been made you will receive an email stating that you have been granted access to the data or explaining the reasons for the rejection of your application.

Reviewing Available Data Through CRN Cloud as an Approved User

Additional tables and functionality become available in ASAP CRN Cloud after you have been approved as a user. The instructions in the section assume that access has been granted.



Note: These instructions assume that you have been granted access to ASAP CRN Cloud data.

Login

To login to ASAP CRN Cloud, select the **Log In** button from the top of the ASAP CRN Cloud or the **Log in with Passport** button from the top of a collection's overview tab. The application will open a Google sign in page prompting you to sign in using the Google-enabled email address you used when submitting your application. Once you have logged in to Google the ASAP CRN Cloud software will re-load the page you navigated away from during the log in process.

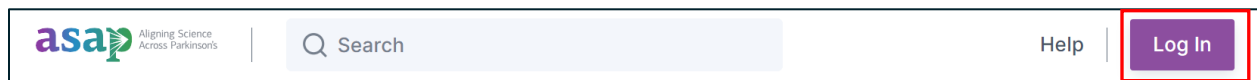


Figure 11- Log in Button

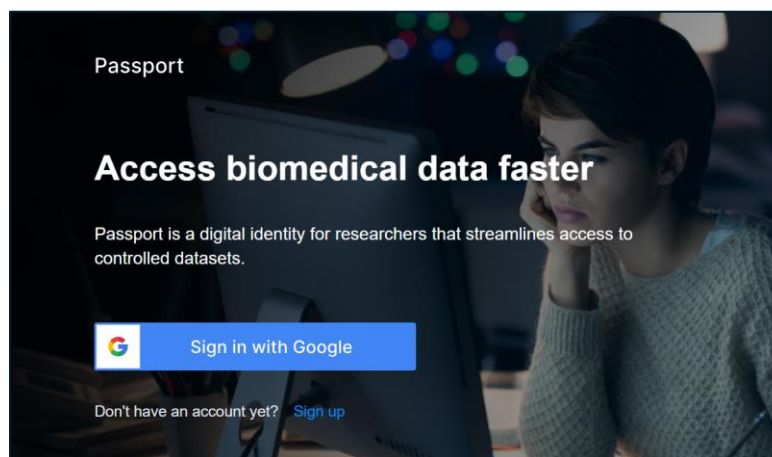


Figure 12- Google Sign in

Review Tables



Note: If using contributed data in a publication, the publication must include the contributing team(s) in the authorship. See the [Authorship List](#) in the Appendix for team details.

Follow the instructions below to review the tables available in member and harmonized collections:

1. Login to ASAP CRN Cloud.
2. Navigate to a member or harmonized collection.
3. Review a table of interest by selecting the tab from the top of the collection or the row of interest from the table at the bottom of the collection's overview tab. When a table is selected the software will display the selected table for review.

Postmortem-Derived Brain Sequencing Collection

| Overview | Clinpath | File Metadata | Protocol | Raw File Metadata | Sample | Study | Subject |
|-------------------|-------------|---------------|-----------------------|----------------------|--------|-------|---------|
| Table Name | Description | Table Size | Created Date | Last Updated | | | |
| Clinpath | | 87 rows | Aug 23, 2023 1:36 PM | Sep 23, 2023 7:20 PM | | | |
| File Metadata | | 126 rows | Sep 14, 2023 10:24 PM | Sep 22, 2023 4:56 PM | | | |
| Protocol | | 2 rows | Aug 23, 2023 1:36 PM | Sep 20, 2023 1:58 PM | | | |
| Raw File Metadata | | 719 rows | Sep 21, 2023 11:50 AM | Sep 22, 2023 4:40 PM | | | |
| Sample | | 186 rows | Aug 23, 2023 1:36 PM | Sep 23, 2023 7:38 PM | | | |
| Study | | 2 rows | Aug 23, 2023 1:36 PM | Sep 23, 2023 7:20 PM | | | |
| Subject | | 37 rows | Aug 23, 2023 1:36 PM | Sep 20, 2023 2:00 PM | | | |

Rows per page: 50

1 - 7

Figure 13 - Selecting a Table for Review

asap | Search | Help | KT

Postmortem-Derived Brain Sequencing Collection

Overview | Clinpath | File Metadata | Protocol | Raw File Metadata | **Sample** | Study | Subject

Search filters

- Adjustment +
- Annotation +
- Assay +
- Assay Ontology Term ID +
- Batch +
- Brain Region +

Results 186 rows

Columns View Query Download

| <input type="checkbox"/> | Omic | Team | Time | Assay | Batch | Dv 200 | Header |
|--------------------------|------|------|------|-------|---------|--------|--------|
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |

Rows per page: 1000 1 - 186

Figure 14 - Review Data

View Query

Follow the instructions below to review the query used to filter the data to the desired subset:

1. Follow the instructions in [Review Tables](#)
2. Select the **View Query** button from within the table review. The software will display a modal containing the query used to populate the table being reviewed.

Postmortem-Derived Brain Sequencing Collection

Overview | Clinpath | File Metadata | Protocol | Raw File Metadata | **Sample** | Study | Subject

Search filters

- Adjustment +
- Annotation +
- Assay +
- Assay Ontology Term ID +
- Batch +
- Brain Region +

Results 186 rows

Columns View Query Download

| <input type="checkbox"/> | Omic | Team | Time | Assay | Batch | Dv 200 | Header |
|--------------------------|------|------|------|-------|---------|--------|--------|
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |

Rows per page: 1000 1 - 186

Figure 15 - View Query

SQL Query

Command Line

```

1  SELECT
2  *
3  FROM "collections"."postmortem_derived_brain_sequencing_collection"."sample"
4  OFFSET 0
5  LIMIT 2000

```

Close

Copy to Clipboard

Figure 16 - Script for Query



Note: Follow the directions within the “How to run an Explorer query using the DNASTack CLI on Verily Workbench” section of the [Explorer Quick Reference Guide](#) to utilize the queries found in Explorer within Verily Workbench.

Download Data

Follow the instructions below to review the query used to filter the data to the desired subset:

1. Follow the instructions in [Review Tables](#)
2. Select the **Download** button from within the table review. The software will download the data in the .CSV format to your local computer’s Downloads folder.

asap

Aligning Science Across Parkinson's

Search

Help

KT

Postmortem-Derived Brain Sequencing Collection

Overview

Clinpath

File Metadata

Protocol

Raw File Metadata

Sample

Study

Subject

Search filters

Adjustment

+

Annotation

+

Assay

+

Results

186 rows

Columns

View Query

Download

| <input type="checkbox"/> | Omic | Team | Time | Assay | Batch | Dv 200 | Header |
|--------------------------|------|------|------|-------|---------|--------|--------|
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |
| <input type="checkbox"/> | RNA | Lee | 0 | v1 | BATCH_4 | NaN | Nan |

Figure 17 - Download Data

Analyzing Data

Accessing Verily Workbench

Reference Workspace

A reference workspace has been set up within Verily Workbench to help users get started. The reference workspace can be duplicated to create a personal workspace that is connected to the ASAP CRN Cloud data for easy analysis using the development environment available in the personal workspace.

To access the Verily Workbench reference workspace, follow the directions below:

1. Navigate to a member of harmonized collection (see [Navigating Collections](#)).
2. Scroll to the bottom half of the collection's Overview tab.
3. Select the link to the Reference Workspace.

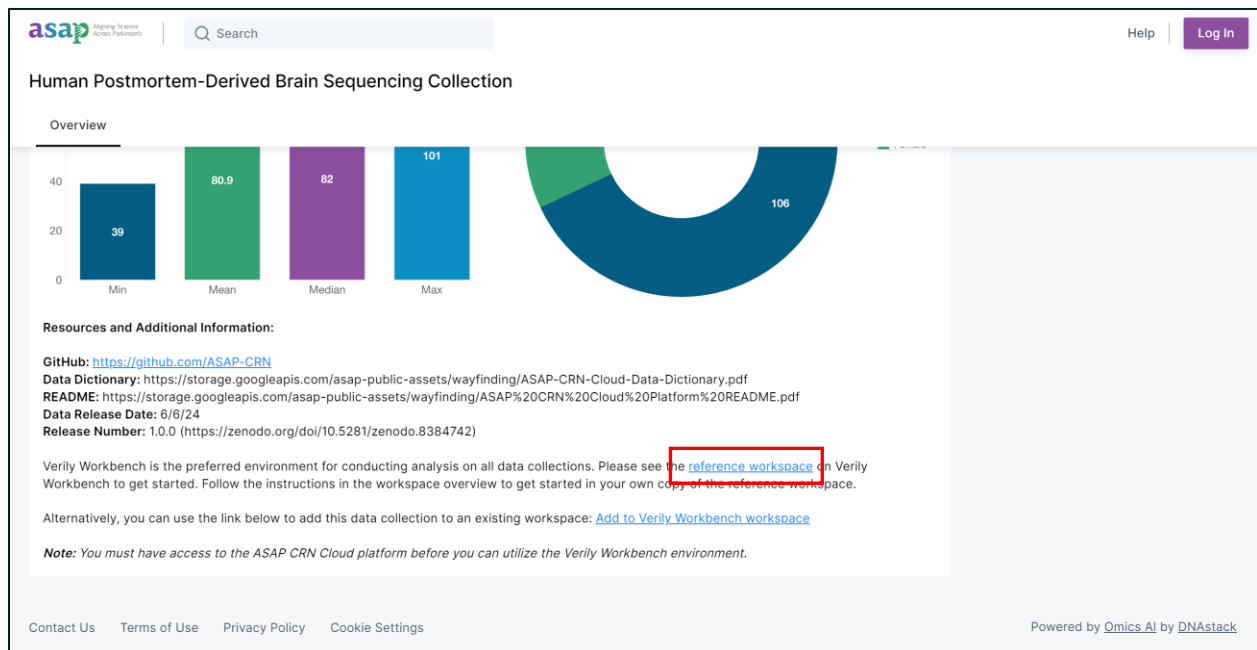


Figure 18 - Link to Reference Workspace

Verily Workbench Setup

Establishing a Personal Workspace

A personal workspace must be established to allow for analysis of data.

To establish a personal workspace, select the Duplicate button from the reference workspace.

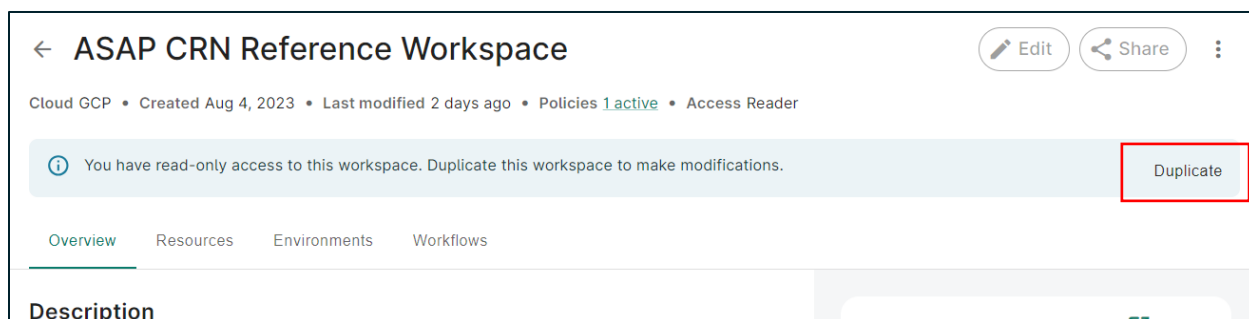


Figure 19 - Duplicate Reference Workspace

Establishing a Spend Profile

General CRN Members and Non-CRN Member Spend Profiles

Please email workbench-support@verily.com to establish your spend profile.

Provisioned CRN Member Spend Pod

A subset of CRN members have been allocated to the CRN pod. If you have received notification that you are within this group, look within the Workspace details section on the right-hand side of your personal workspace. Confirm that you are correctly assigned to the ASAP CRN Cloud pod (asap-crn-gcp).

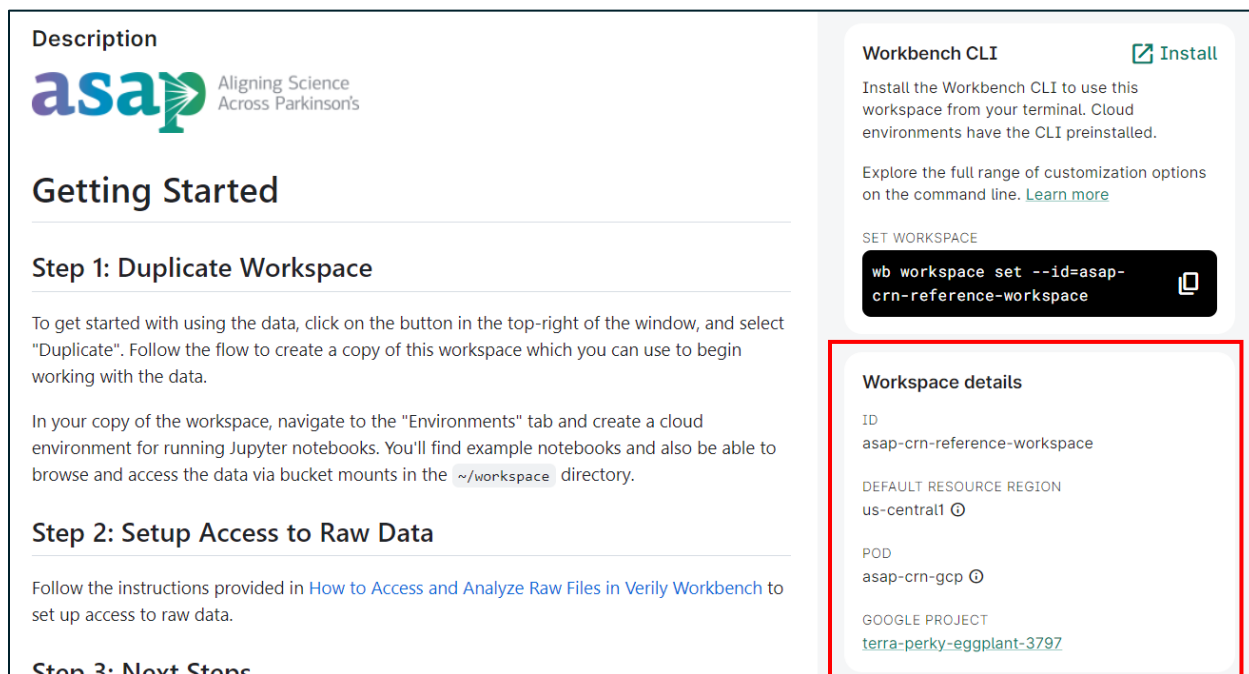


Figure 20 - Locating Workspace Details

Workspace details

ID
asap-crn-reference-workspace

DEFAULT RESOURCE REGION
us-central1 ⓘ

POD
asap-crn-gcp ⓘ

GOOGLE PROJECT
[terra-perky-eggplant-3797](#)

Figure 21 - Confirm Spend Profile



Note: Please use the spend profile responsibly and exclusively for projects directly related to ASAP. Responsible spending will help the ASAP CRN program continue to support cloud costs. Please follow the best practices outlined in Verily Workbench's [Cloud Cost Management](#) guide.

Establishing a Development Environment

After establishing a personal workspace, follow the instructions for setting up a development environment found here: [Explorer Programmatic Access](#).



Note: The instructions within the Verily Workbench reference workspace will not be available until your access has been established. (See [Submitting for Data Access](#))



Note: Ensure that you have selected **us-central1** as your Default Resource Region to avoid extra cloud costs.

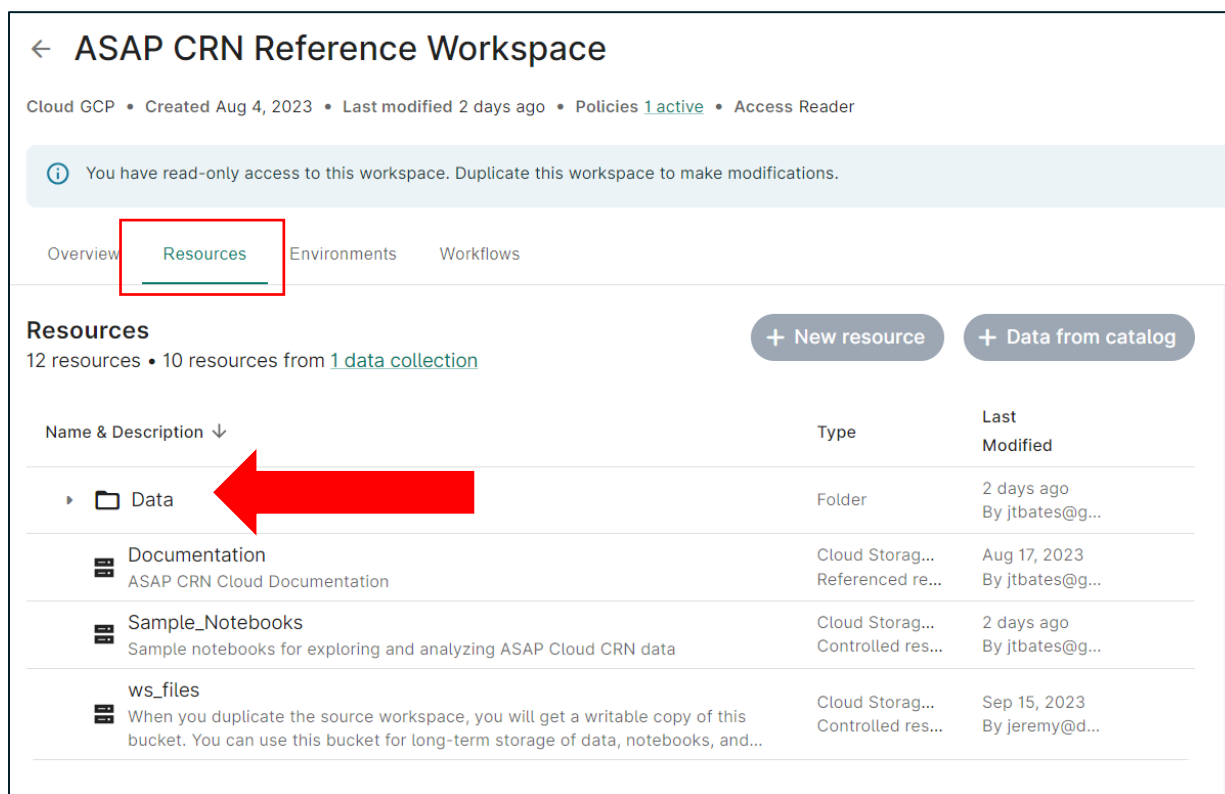
Accessing Data

Navigating Data on Verily Workbench

The data available through ASAP CRN Cloud includes raw data (e.g. – FASTQs) and curated data (processed data derived from raw data).

Curated Data

To review curated data, navigate to your personal workspace then select the Resources tab.



← ASAP CRN Reference Workspace

Cloud GCP • Created Aug 4, 2023 • Last modified 2 days ago • Policies [1 active](#) • Access Reader

📘 You have read-only access to this workspace. Duplicate this workspace to make modifications.

Overview **Resources** Environments Workflows

Resources
12 resources • 10 resources from [1 data collection](#)

+ New resource + Data from catalog

| Name & Description ↓ | Type | Last Modified |
|--|--------------------------------------|---------------------------------|
| ▶ Data | Folder | 2 days ago By jtbates@g... |
| Documentation ASAP CRN Cloud Documentation | Cloud Storag... Referenced re... | Aug 17, 2023 By jtbates@g... |
| Sample_Notebooks Sample notebooks for exploring and analyzing ASAP Cloud CRN data | Cloud Storag... Controlled res... | 2 days ago By jtbates@g... |
| ws_files When you duplicate the source workspace, you will get a writable copy of this bucket. You can use this bucket for long-term storage of data, notebooks, and... | Cloud Storag... Controlled res... | Sep 15, 2023 By jeremy@d... |

Figure 22 - Curated Data

Raw Data

Raw data is available in **Requester Pays** buckets, meaning that the requestor (i.e., you) must assume the cost for data transfer when requesting raw data. ASAP CRN Cloud requires users to establish a Google Cloud Billing Account owned by you (i.e., the same Google Identity you use to log in to CRN Cloud) to access the raw data.

To keep costs minimal, please follow the best practices provided in Verily Workbench's [Cloud Cost Management](#) guide.

Transfer Raw Data from Requester Pays Buckets

To setup a Google Cloud Billing Account and access the raw data, follow the instructions found here: [Setup Requester Pays Bucket and Access Raw Data](#).

Loading Data Using Saved Searches

Signed URLs

Signed URLs can be used to save custom search results, established through the CRN Cloud website, into an associated data bucket.



Note: The “Run the sample Jupyter notebook” instructions are found within the Next Steps section. You may need to scroll down to the lower half of the Overview tab.

Follow the instructions below to utilize the signed URLs functionality:

1. Navigate to cloud.parkinsonsroadmap.org and log in.
2. Open collection card of interest.
3. From within the collection card, use the search and selection features to view desired data. (See [Collection Search and Programmatic Access](#) for more details on CRN Cloud’s search functionality.)

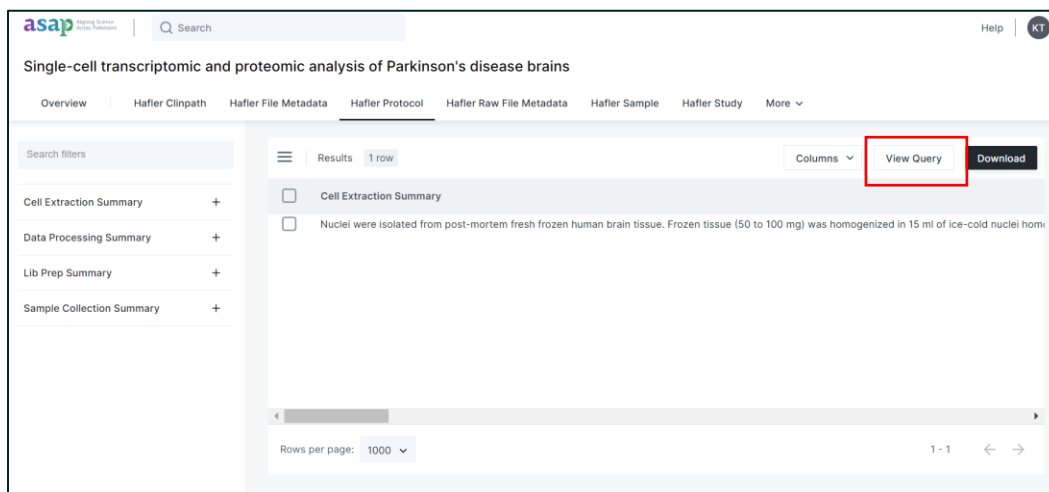


Figure 23 - Select View Query

4. Select the **View Query** button, then the **Signed URL** tab to access the signed URL.

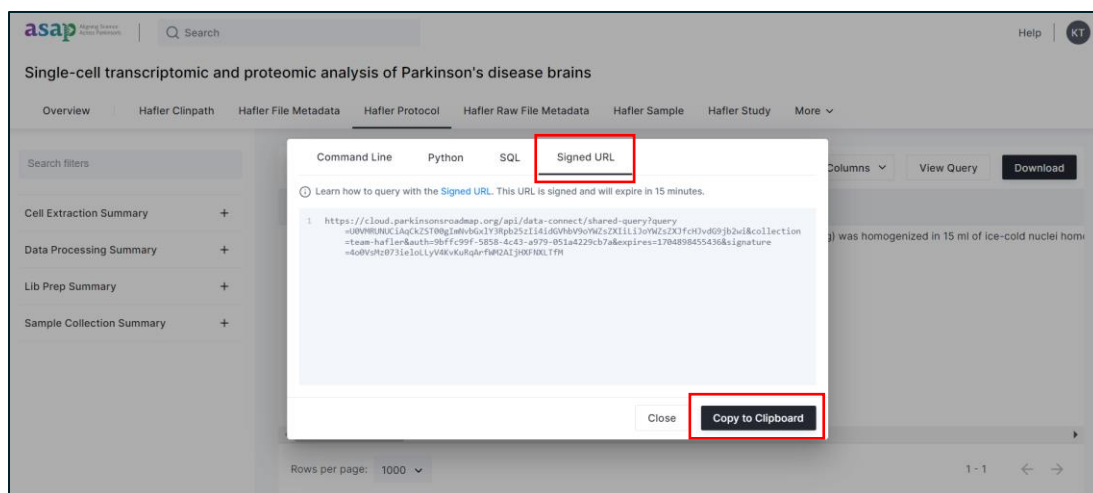



Figure 24 - Copy Signed URL

5. Select the **Copy to Clipboard** button.



Note: Once copied, the Signed URL is only valid for 15 minutes. You will have to refresh and copy the signed URL again if the URL has expired.

6. Navigate to [Verily](https://workbench.verily.com) (workbench.verily.com) and open your workspace.
7. Navigate to the **Resources** tab, then to an existing cloud bucket in the workspace.

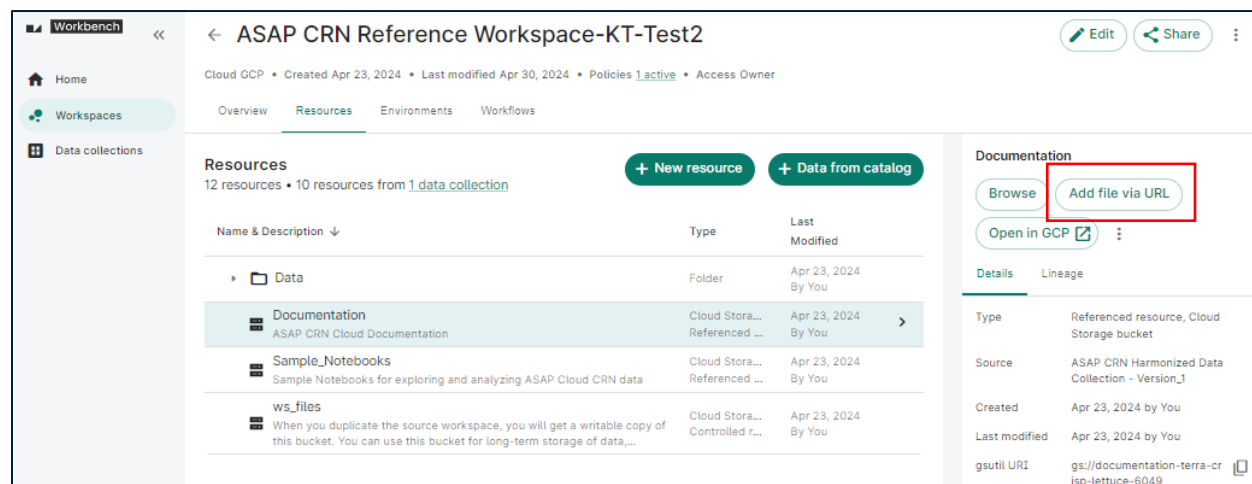


Figure 25 – Add file via URL

8. Select **+ Add file via URL**. In the resulting dialog window:
 - a. Enter the **Signed URL** from into the **URL** section.

- b. Choose a file name for the resulting file.
- c. Click **+ Add file to bucket** to save the search query results to a bucket.
Note: Data will be saved as a CSV.
- d. You may click **Preview** to browse the data in the Verily Workbench UI.
The data is also available for any analyses/notebooks.



Note: For additional guidance, follow the instructions provided by [Verily's Add Data Via URL](#).

Python code snippets for Jupyter Notebooks

Python code snippets can be added to a Jupyter notebook to load custom search results, established through the CRN Cloud website, into the associated notebook.

Follow the instructions below to utilize Python code snippet functionality:

1. Navigate to cloud.parkinsonsroadmap.org and log in.
2. Open collection card of interest.
3. From within the collection card, use the search and selection features to view desired data. (See [Collection Search and Programmatic Access](#) for more details on CRN Cloud's search functionality.)
4. Select the **View Query** button, then the **Python** tab to access the associated Python script.
5. Select the **Copy to Clipboard** button.
6. Navigate to [Verily](https://workbench.verily.com) (workbench.verily.com) and open the desired workspace.

(Next steps assume you already have a Jupyter environment running in VWB.)

7. In Verily Workbench, navigate to a running Jupyter Lab environment.
8. Start a new Jupyter notebook.
9. Install the dnastack-client-library by running: `pip install -U dnastack-client-library`
10. Paste the code snippet copied in Step 5 into the Jupyter notebook and run. You will be prompted to login to the ASAP CRN Cloud via a web URL. Paste the URL into a web browser and follow the prompts to sign in. Once the web page displays "you may return to your device", you may return to the Jupyter Notebook.
11. The search results are loaded from the object **result_iterator**. For example, you can unpack the results in a **for loop**:

```
for result in result_iterator:
    print(result)
```

Analytical Functionality

To analyze the curated or raw data, utilize the development environment established in [Establishing a Development Environment](#).

To navigate to an existing development environment, select the Environments tab within your personal workspace. Open the desired cloud environment by double-clicking the name of the environment (in green text).

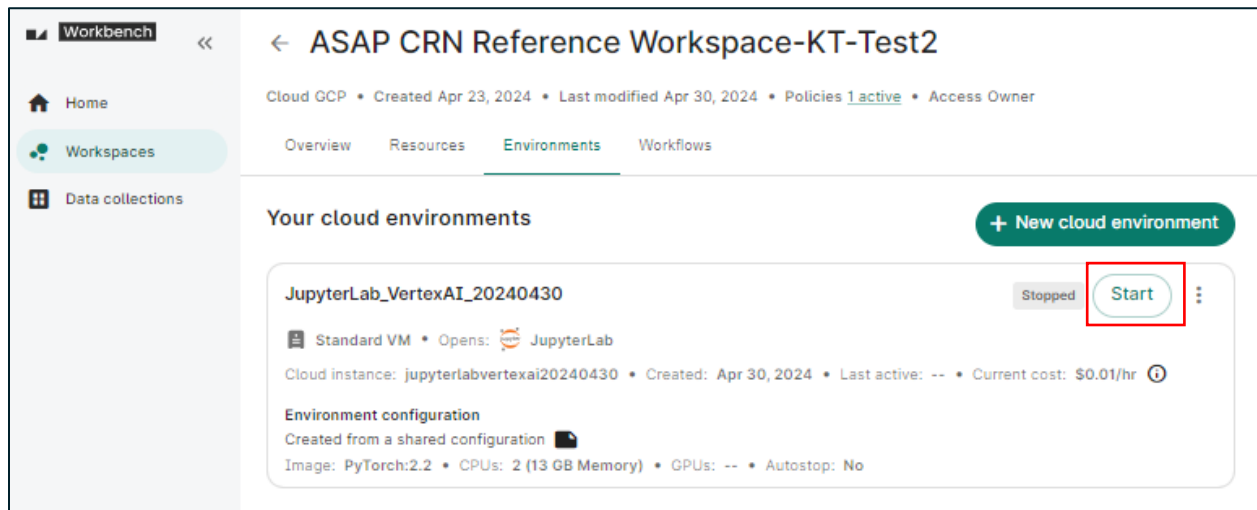


Figure 26 - Open an Existing Development Environment



Figure 27 - Double Click Name of Dev Environment to Open

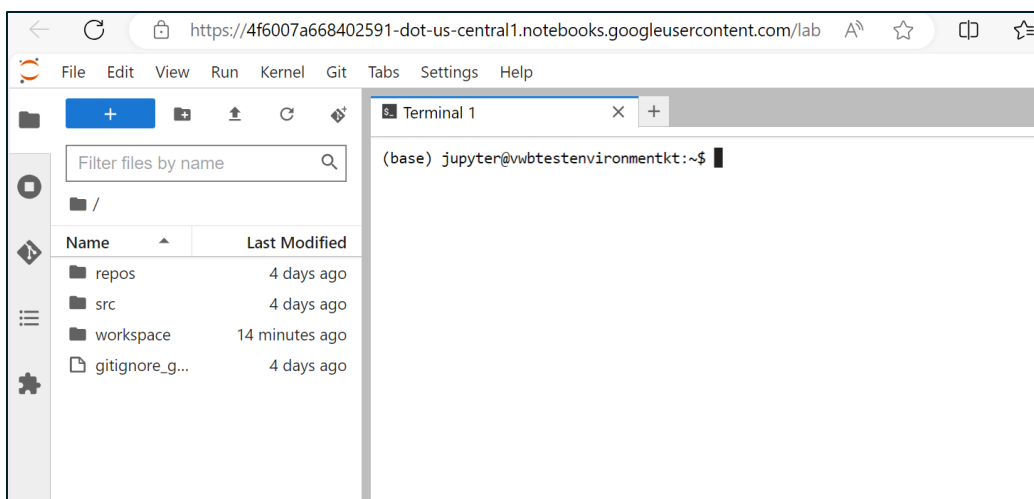


Figure 28 - Development Environment

Appendix

Submitting Data to CRN Cloud

CRN Cloud data submission is open to all members of the ASAP Collaborative Research Network. If you're ready to begin uploading your data, please follow the steps below for each dataset being contributed:

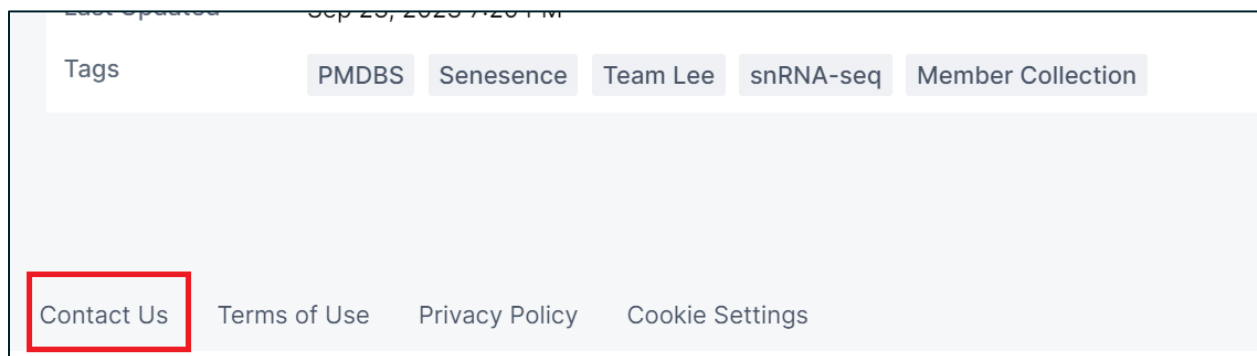
1. Select a member of your team to act as the point of contact.
2. Have the point of contact complete an [ASAP CRN Data Curation Lab Survey](#) for each dataset they contribute.
3. Meet with the CRN Cloud data curation team to discuss the data being contributed.
 - a. Use the following [booking page](#) to set up a time to meet with our data curation team.
4. Each point of contact contributing data submits metadata and follows the provided instructions to upload data.
 - a. Detailed explanation and reference for the [metadata assembly](#),
 - b. Data transfer [instructions](#)

Data Dictionary

The data dictionary for ASAP CRN Cloud can be found here: [Data Dictionary](#).

Help and Troubleshooting

For any software issues, please contact DNASTack by selecting the Contact Us link in the bottom left corner of the screen (within ASAP CRN Cloud). This will open an email addressed to support@dnastack.org.



If you need additional support, please contact cloud@parkinsonsroadmap.org.

Resources for Contributing Teams

Resources for contributing teams can be found on the ASAP Hub within the CRN Cloud Working Group in a folder named [Data Use Review \(Data Contributors Only\)](#).

Contributing teams can find information explaining the intended use of their data within this folder. If you are a contributing team member and would like help with reviewing the intended use of your data, please contact cloud@parkinsonsroadmap.org.

Authorship List

The Authorship List provides an easy to reference list of contributors to the ASAP CRN Cloud data. It is intended to be used to provide credit to contributing teams when publishing data produced using ASAP CRN Cloud data.

To use the [Authorship List](#), use the link to navigate to the document, locate the name of the data set used in your data within the document, and copy the listed names into the appropriate acknowledgements section in your paper.