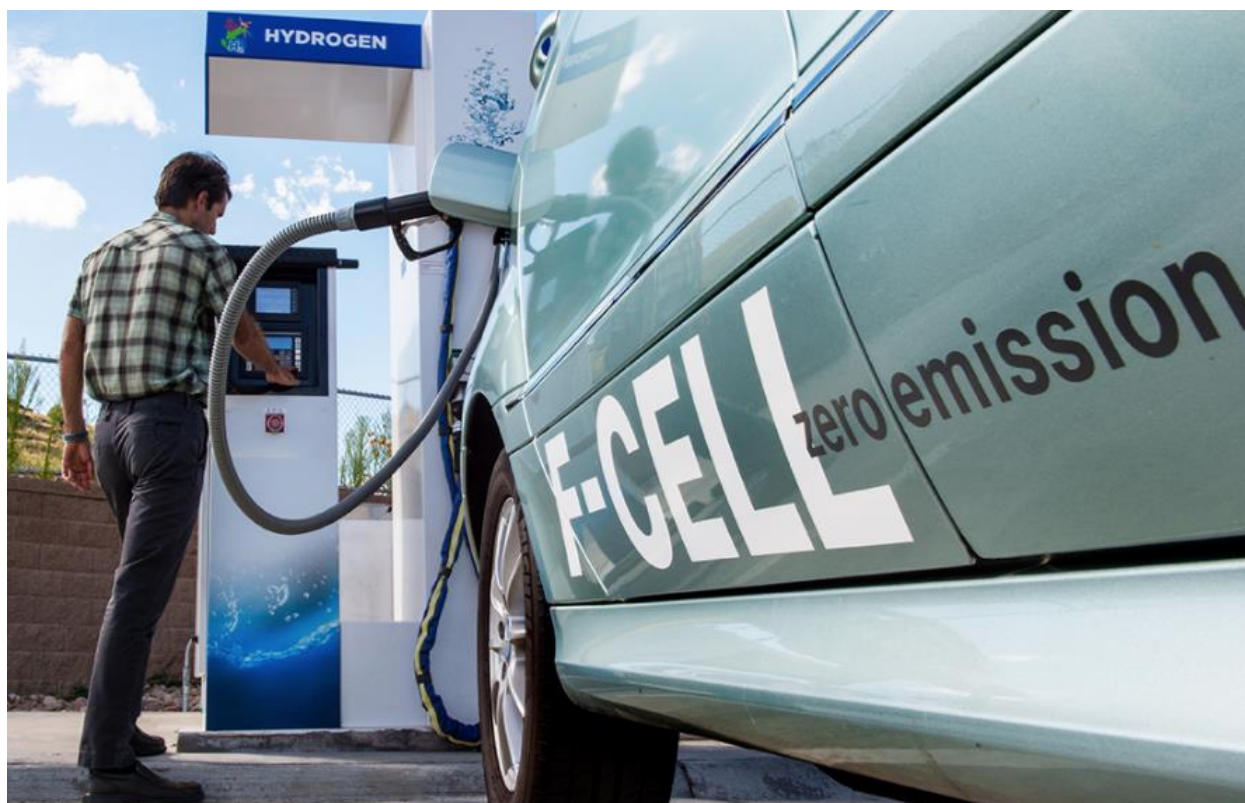


Using the ElectroCat Data Hub



U.S. DEPARTMENT OF
ENERGY

Energy Efficiency &
Renewable Energy



Energy Materials Network
U.S. Department of Energy

November 20, 2017

1

Table of Contents

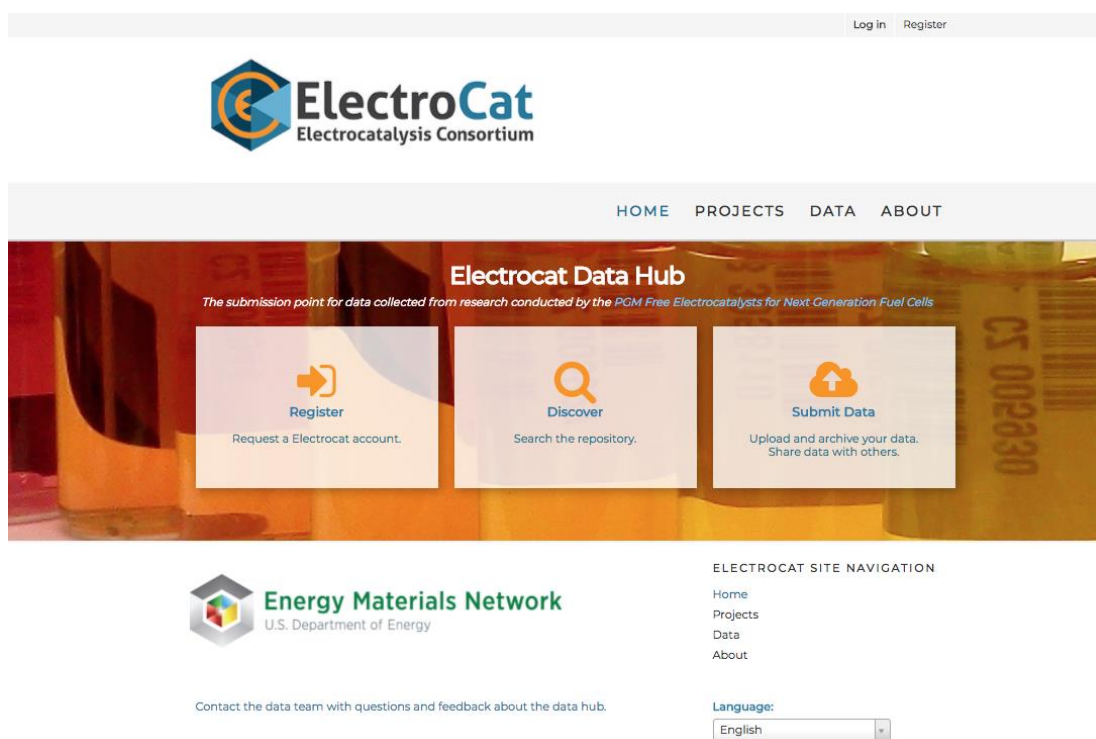
Using the ChemCatBio Data Hub	1
Overview	3
Registration, Login and Security	3
Projects & Data	5
Datasets	6
Files & Resources	6
Metadata	6
Search	6
For All Researchers	7
Adding Data	7
Adding Data to an Existing Dataset	11
Deleting a File or Resource	13
Viewing Data	14
Accessing the Data through the API	17
For Project Principle Investigators (PIs).....	18
Adding new members to a project	18
Changing permissions for a member of a project	18
Removing a project member	18
Questions and Feedback?	18



Overview

The ElectroCat Data Hub is a platform for consortium members and partners to share data and ideas. Researchers are encouraged to place their data within this hub, where it can be protected and distributed as needed. Providing data to the hub can increase communication efficiency between all parties and create a seamless environment for eventual releasing of data from DOE-funded research. Data can be compartmentalized and secured by project or scope and is able to store a wide variety of data types and files. The hub allows for searching the metadata and data of all resources stored within, providing a method to support discovery.

The hub is built around the [Comprehensive Knowledge Archive Network](#) or CKAN software framework, but has been extended through plugins and code rewrites beyond the original distribution package. The focus of the data hub is to provide the users with a system that can provide the efficiency and security for collaborative data sharing and public release of data, as specified by DOE data requirements.



The architecture of the Data Hub can be divided into three main areas

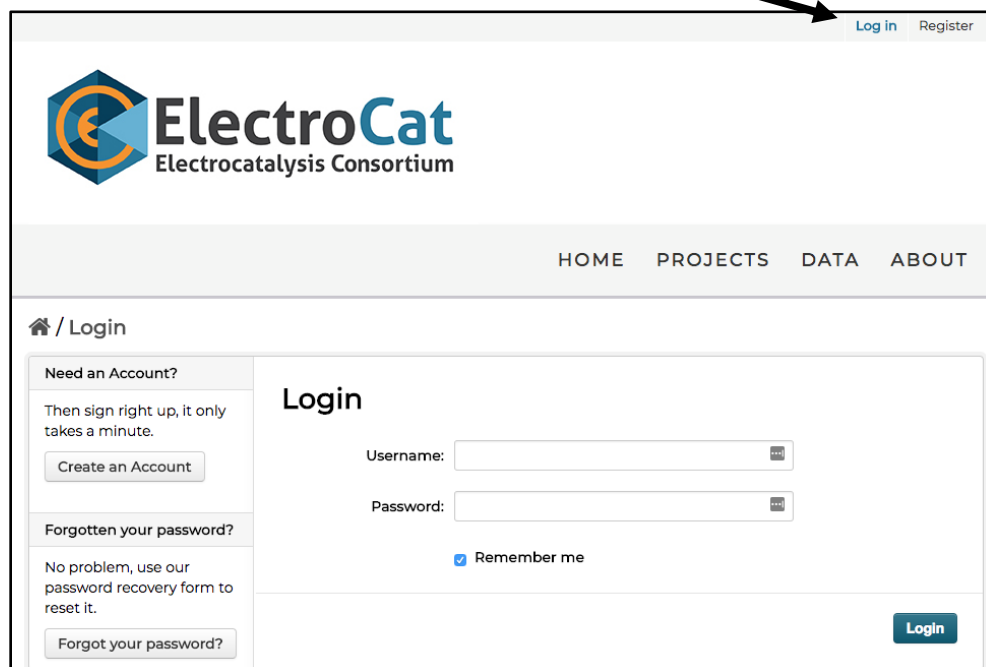
- Registration, Login and Security
- Projects & Data
- Search

Registration, Login and Security

You must be a registered user of the data hub and a consortium member to be able to access any files that are not deemed **"Public"**. The registration process is simple and begins by clicking the large "Registration" button on the main page (see image above), which will take user to the registration page.

1. Enter the required information on the registration page and click **Create Account**.
 - a. **National Lab users** should use their lab username and lab email address.

2. The researcher needs to **email the administrator** (emnadmin@nrel.gov) with the following details:
 - a. The Institution you work for.
 - b. The username you registered as.
 - c. The list of **projects you need access to**. *You can review the list of current projects without registration by clicking on the Projects tab on the menu bar at the top of the page.*
3. The administrator will contact the Project PI for **access approval**.
4. Once given access to the site you may **Login** to the data hub.

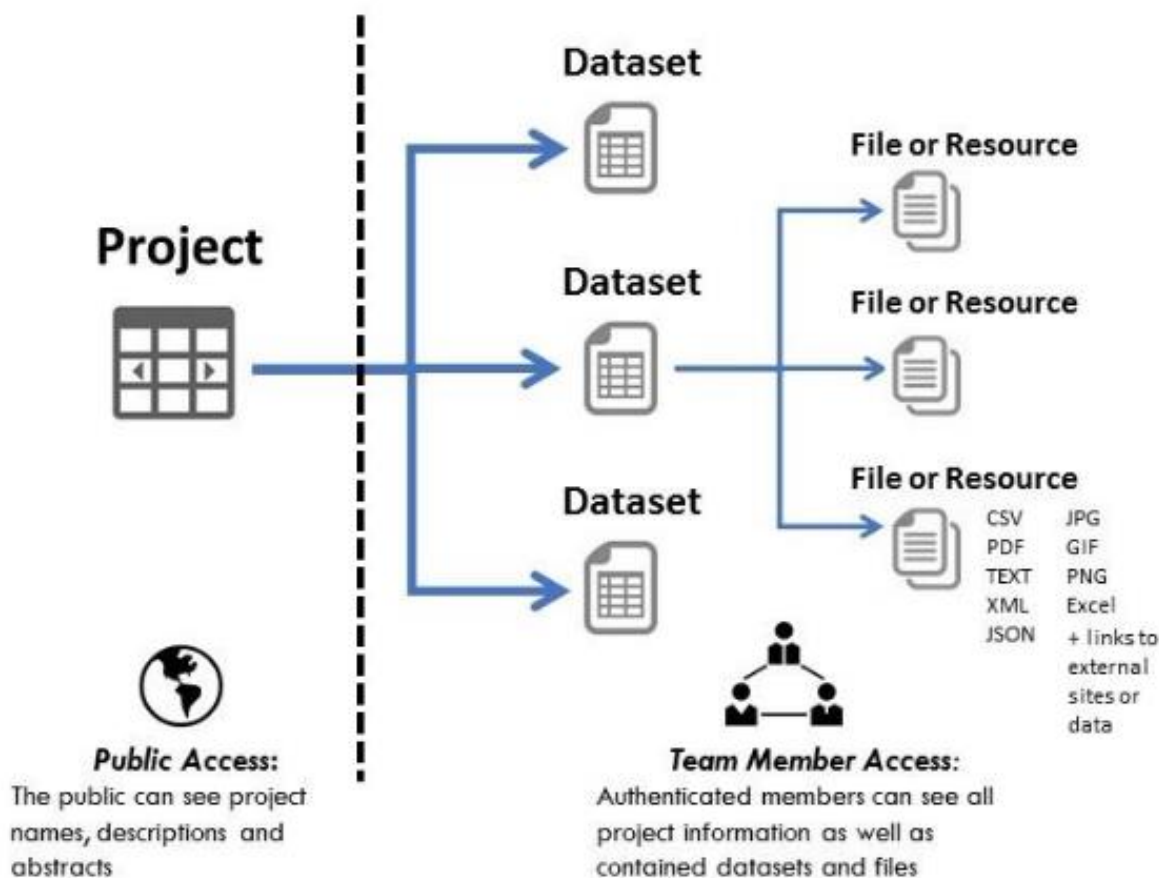


The screenshot shows the ElectroCat website interface. At the top right, there are links for 'Log in' and 'Register'. Below the header is a navigation menu with 'HOME', 'PROJECTS', 'DATA', and 'ABOUT'. The main content area is titled '/ Login'. On the left, there are links for 'Need an Account?' and 'Forgotten your password?'. The central 'Login' form includes fields for 'Username:' and 'Password:', a 'Remember me' checkbox, and a 'Login' button.

Once given access to the site, you can request additional projects by contacting the administrator (emnadmin@nrel.gov).

By default, all datasets created and files uploaded are designated as **Private**, meaning only members of that project can access them. In time, there will be a formal process implemented for converting the data to other access levels such as embargoed and public.

Unregistered or public users can see basic details about any project and can view and download any designated public data. However, any private datasets cannot be seen or accessed by the public.



Projects & Data

The awarded projects within the ElectroCat Consortium become the Projects on the Data Hub. Projects on the Data Hub are then used to organize data further by creation and upload of distinct Datasets and Resources (see below).

The “User Resources” project showcases ElectroCat specific data tools and user guides for the Data Hub. All consortium members are given access to the “User Resources” project so they may review the custom data tools and Data Hub user documents.

All projects and sub-projects are created by a system administrator. Contact the [administrator](#) for any projects or sub-projects you need created or adjusted.

A project can contain any number of datasets (see below) and sub-projects. A sub-project allows for additional granularity for storing data within a project. Sub-projects also allow for restricting data within a project to only a sub set of the overall project members.

Datasets

From the Home page, the middle large button, “**Discover**”, will take you directly to the Datasets you have access to. If you are not logged in, you will only see Public Datasets. A dataset is similar to a folder in most computer systems. Datasets are used to organize data into logical areas and each Dataset can contain any number of files or resource links.

By default each Dataset is established as private, giving only project members access to the data. At a future date, we will set up procedures for public release of data. If you have a current dataset that needs to be made public, please contact the [administrator](#) and we will help you.

The third large button, “**Submit Data**” will prompt you to login with the credentials you established at the time you registered on the Data Hub. Once you are logged in, you can either navigate to the Project to which you have access to “Add New Resource” or click on the “Submit Data” button from the Home tab to directly create a Dataset under the appropriate Project to which you have access (See “[Adding Data](#)”).

Files & Resources

All data uploaded to the hub will be either a file or a web link (URL). Collectively these are called resources. The data hub is capable of housing any type of file format. We have limited the upload size to 5MB/file, but if the situation arises that larger files need to be stored, please contact the [administrator](#).

Metadata

Researchers can create datasets to upload files and create links to external public tools and public resources (contact the [administrator](#) if you have a public tool that needs to be made public). Metadata (data that describes the data) is a set of information that describes the file being uploaded. Datasets have associated metadata which the researcher is prompted to enter upon Dataset creation. Dataset metadata includes Institution, Author, Maintainer E-mail, Sample Barcode, Collection Date, Data Source Type (External Data or Tools, Historical or Literature, Lab Experimental, Modeling and Simulation). Each user uploading data into the hub will need to fill out any associated metadata. Eventually, there will be resource metadata, which will allow for improved search capabilities within projects.

Search

The Data Hub allows you to search on different criteria defined within the metadata, user-defined Tags, as well as within the description of the Project, Dataset, or Resource. You can search for data from the “**Discover**” button on the Home tab and from the **search bar** on the Projects or Data tab. You can search on the file type, to see all CSV files, for example; you can search on keyword or data source type (External Data or Tools, Historical or Literature, Lab Experimental, Modeling and Simulation). The Data tab also displays the left-hand **metadata faceted search** options. The faceted search shows pieces of metadata that have been identified at the Dataset or Resource level with a parenthetical reference of the number of times that metadata is being used. **Your results will only return data you have access to.**

The screenshot displays the Energy Materials Network Data Hub interface. On the left, a sidebar contains navigation links for Projects, Tags, Institutions, and Data Source Types. The main content area shows search results for datasets, including 'Greenway Nano' and 'CMU Cathode'. A blue arrow labeled 'Metadata Faceted Search' points to the sidebar. Overlaid on the right is a 'Create Dataset' form with fields for Title, Project, Description, Tags, and Dataset Metadata. A blue arrow labeled 'User-defined Tags entered upon Dataset and Resource creation' points to the Tags field in the form.

Metadata Faceted Search

User-defined Tags entered upon Dataset and Resource creation

For All Researchers

Adding Data

The key to the data hub is uploading research data that should be shared with project members and eventually to the public. The process may require the creation of a new dataset or it could be adding new data files to an existing project's dataset. During the creation of the dataset and / or adding a resource or file, you will be prompted to provide additional information (metadata) that can facilitate understanding, searching, and organizing the data.

Home / Projects

Projects

All Projects

My Projects

Administration

Add Project

Project Tree

- CMU Cathode
- Giner Mn Catalysts
- Greenway Nano
- User Resources
- PNNL Electrocatalysts
- Thin Film Model System Core Capability

Search projects...

Order by: Name Ascending

CMU Cathode
 0 Datasets
 Advanced PGM-free Cathode Engineering for High Power Density and Durability...

Giner Mn Catalysts
 0 Datasets
 Durable Mn-based PGM-Free Catalysts for Polymer Electrolyte Membrane Fuel Cells...

Greenway Nano
 0 Datasets
 PGM-free Engineered Framework Nano-Structure Catalysts Recipient Greenway...

User Resources
 1 Dataset
 A collection of resources for registered data hub users. Examples of site...

PNNL Electrocatalysts
 0 Datasets
 Highly Active and Durable PGM-free ORR Electrocatalysts through the Synergy...

Thin Film Model System Core Capability
 1 Dataset
 Thin-film synthesis and characterization of PGM-free catalysts Abstract This...

2. Click on the Project that will store the data

Home / Projects / Giner Mn Catalysts

Project

Overview

Datasets

Activities

Administration

Edit Project

Bulk Operations

Add Dataset

Members

Project Tree

- Giner Mn Catalysts

Giner Mn Catalysts

Durable Mn-based PGM-Free Catalysts for Polymer Electrolyte Membrane Fuel Cells

Recipient Giner, Inc (PI: Xu, Hui)

Subs Prof. Gang Wu (University at Buffalo, the State University of New York), Prof. Guofeng Wang (University of Pittsburgh), Dr. Anusorn Kongkanand (General Motors Company)

Abstract Polymer electrolyte membrane fuel cell (PEMFC) is a technology to generate electricity through efficient electrochemical conversion of hydrogen and oxygen into environmentally benign water. Hence, it is of great economic and societal benefits to apply PEMFC in automotive

1. Click on **Add Dataset**, left panel

1 Create dataset **2 Add data**

Title: eg. A descriptive title ⓘ
* URL: datahub.electrocat.org/dataset/<dataset> Edit

Project: Giner Mn Catalysts ▼

Description: eg. Some useful notes about the data
You can use Markdown formatting here

Tags: eg. economy, mental health, government

Dataset Metadata

* Institution: -- select an option -- ▼

* Author: admin ▼

* Maintainer Email:

Sample Barcode:

Collection Date: 4/17/2018

Data Source Type: -- select an option -- ▼

Comments:

The data license you select above only applies to the contents of any resource files that you add to this dataset. By submitting this form, you agree to release the metadata values that you enter into the form under the Open Database License.

* Required field

Next: Add Data

3. Complete all Fields on form

Verify that the set *Project* is correct, you could place the dataset in another project, but make sure you have access to that target project.

For Tags use single words if possible. You need to hit **Return** after each tag for it to appear.

4. Click on **Add Data** to move to next page

5. Choose to upload a file or create a link.

Clicking the **Upload** button will open a dialog box for you to choose a file from your computer to upload. The **Link** button will ask you to provide a URL.

6. Complete all Fields on the form.

Cost Model / Add New Resource

1 Create dataset **2 Add data**

Data:

Name: eg. January 2011 Gold Prices

Description: Some useful notes about the data
You can use Markdown formatting here

Format: eg. CSV, XML or JSON ▼
 ⓘ This will be guessed automatically. Leave blank if you wish

Previous **Save & add another** **Finish**

7. Click **Save & add another** to open a new form for another resource or click **Finish** to complete the upload.

8. Dataset creation and resource upload complete



Once everything has been completed using the Finish button, a page for the dataset will appear showing all resources currently within it.

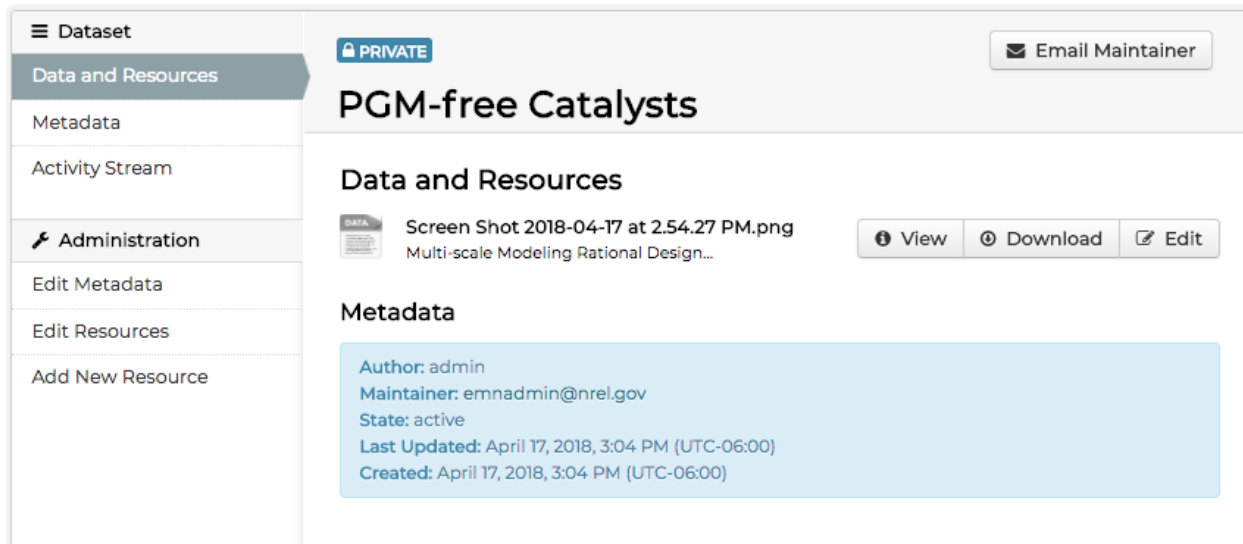
The navigation buttons on the left panel:

Edit Metadata - Edit the Dataset's current metadata.

Edit Resources - Opens the resource list for the dataset. From there, you can add a new resource and edit the order they appear in the dataset list.

Add New Resource - Opens the resource upload page (as above).

🏠 / Projects / CMU Cathode / PGM-free Catalysts



The buttons to the right of the resources:

View - If it is a “viewable” resource within the data hub, it will be displayed along with the associated metadata. Currently CSV, TXT, most picture formats, XML, and JSON are directly viewable in the hub. Link Resources will open a new tab within your browser to display the web page. Non-viewable resources will not be displayed but their metadata will be.

Download – Will download the resource file through your browser to your local computer.

Edit – Allows you to edit the metadata associated with that file or resource.

Adding Data to an Existing Dataset

Within some working projects, depending on how the datasets are being used, you may need to continue adding data to an existing dataset. Example: A project could have a dataset for all XRD measurements. The dataset metadata could be the *Modeling and Simulation Data Source Type*, and each data file could cover separate measurements for different samples (e.g., prepared by using different conditions), so can show different sample names, synthesis techniques/conditions, and/or pre- and/or post-treatments of a sample.



1. Click on the Project that will store the data

2. Click on **Datasets**, left panel

The screenshot displays the ElectroCat web application interface. On the left, a sidebar menu lists various projects: CMU Cathode, Giner Mn Catalysts, Greenway Nano, User Resources, PNNL Electrocatalysts, and Thin Film Model System Core Capability. A blue arrow points from the '1. Click on the Project that will store the data' instruction to the 'CMU Cathode' project in the sidebar. The main content area shows a list of projects with a search bar and an 'Order by' dropdown. A second blue arrow points from the '2. Click on **Datasets**, left panel' instruction to the 'Datasets' tab in the left panel of the 'CMU Cathode' project view. The 'CMU Cathode' project view shows a list of datasets, with the 'CMU Cathode' dataset selected. The dataset details for 'CMU Cathode' are displayed on the right, including the project name, description, and a list of team members.



Home / Projects / CMU Cathode

Project
Overview
Datasets
Activity Stream
Administration
Edit Project
Bulk Edit Datasets
Add Dataset
Members
Tags
ORR (1)
PEFC (1)
PGM-free (1)

CMU Cathode

Order by: Relevance

PRIVATE PGM-free Catalysts
1 Resource
Multi-scale Modeling Rational Design of PGM-free Catalysts

PRIVATE Auotmotive PEFC
1 Resource
This dataset has no description

3. Click on the dataset to add data to

Home / Projects / CMU Cathode / PGM-free Catalysts

Dataset
Data and Resources
Metadata
Activity Stream
Administration
Edit Metadata
Edit Resources
Add New Resource

PRIVATE
Email Maintainer

PGM-free Catalysts

Data and Resources

Screen Shot 2018-04-17 at 2.54.27 PM.png	View Download Edit
electrolysis-pol-curve.csv	View Download Edit
Shared Team experiment.pdf	View Download Edit

Metadata

Author: admin
Maintainer: emnadmin@nrel.gov
State: active
Last Updated: April 17, 2018, 3:27 PM (UTC-06:00)
Created: April 17, 2018, 3:04 PM (UTC-06:00)

4. Click **Add New Resource**

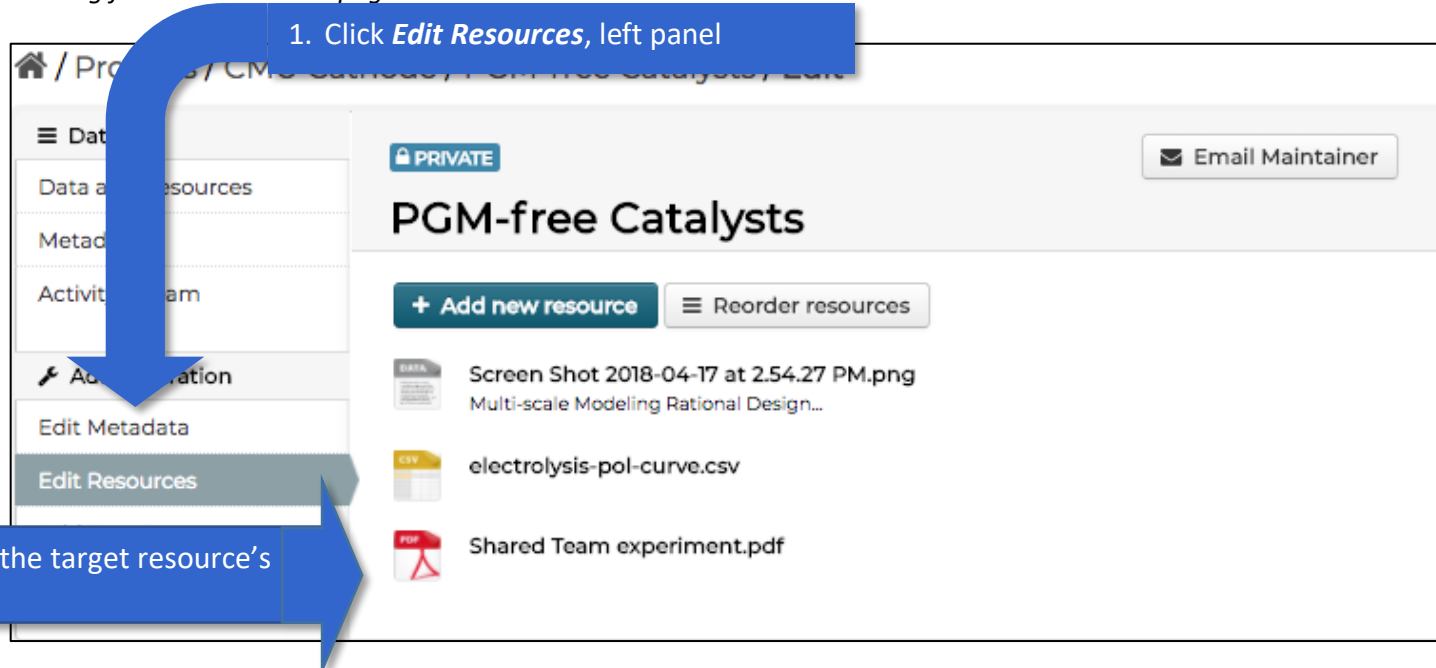
From that point, follow the directions in the [previous section](#), from step 5 and onward.

Deleting a File or Resource

There will be times when a user might need to delete a file or resource that has been added to a dataset. Perhaps you have uploaded the wrong file or noticed a mistake and need to reload the file. Follow the directions below to delete a Resource.

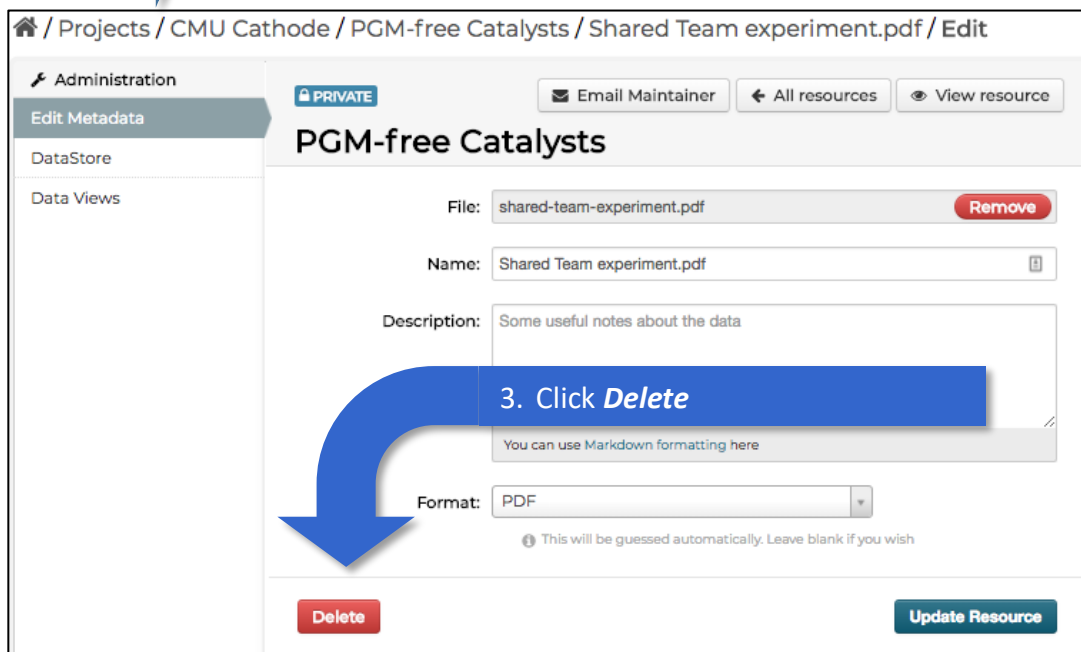
Starting from the Dataset's page

1. Click **Edit Resources**, left panel



2. Click on the target resource's name

3. Click **Delete**



4. A pop up dialog box will then appear and ask you to confirm deleting this file

After you confirm deleting the file you will be brought back to the dataset's resource list page.

Viewing Data

In many cases the ElectroCat data hub has the inherent ability to render the data file for viewing directly in the browser. This ability to view data can also be extended through custom coded plugins. View demonstrations of some of the viewing capabilities in the **User Resources** project on the data hub. To utilize the basic display capability of the system, do the following:

Starting from the Dataset's page

1. Click **View**, to right of the data or resource name.

2. View page opens with the format's display

Home / Projects / User Resources / Data Tools

Dataset

- Data and Resources
- Metadata
- Activity Stream

Administration

- Edit Metadata
- Edit Resources
- Add New Resource

PRIVATE

Data Tools

Data and Resources

- Data Explorer**
This is a demo of the Data Explorer...
[View] [Download] [Edit]
- Map View**
This is a demo of the Map view, which...
[View] [Download] [Edit]
- IMAGE PGM free Catalyst.png**
Image files, such as PNG and GIF,...
[View] [Download] [Edit]

Metadata

Author: admin
Maintainer: emnadmin@nrel.gov
State: active
Last Updated: April 20, 2018, 12:29 PM (UTC-06:00)
Created: November 3, 2017, 3:01 PM (UTC-06:00)

If the format has been predefined it will open that within the viewing engine. Additional view types can be created as plugins. This is an example of an image file type.

IMAGE PGM free Catalyst.png [Email Maintainer] [Manage] [Download]

Image files, such as PNG and GIF, will automatically render as an image. This is an example of how an image is displayed.

[Image]

[Embed]



With **CSV files** the **Data Explorer view** will allow for the data to be viewed in three possible ways: **Table**, **Graph** or **Map**. To utilize the Graph or Map the data has to lend itself to being viewed in that manner and both will require interaction with the user to display the data as needed. The map function requires either Latitude and Longitude or GeoJSON be included in the file for it to function correctly.

Across from the resource name are 4 buttons:

Email Maintainer – Opens an email to the person listed as the maintainer of the dataset.

Manage – This allows you to change the metadata for the resource.

Download – This will download the file through your browser to your local computer.

Data API – This green button will open another dialog box listing the syntax to access this particular data file programmatically ([see below](#)).

Email Maintainer
Manage
Download
Data API

Data Explorer

This is a demo of the Data Explorer tool, which comes with CKAN and is automatically available anytime you upload a CSV file. This tool lets you explore a CSV file using a grid, filter the data, create graphs using any of the columns in the file, and (if you have geo information such as latitude and longitude) you can view a Map of the data points.

To use Data Explorer: Click "Graph" and follow the instructions.

Data Explorer
Embed

Add Filter
Grid
Graph
Map
290 records
1 - 100
Search data ...
Go
Filters

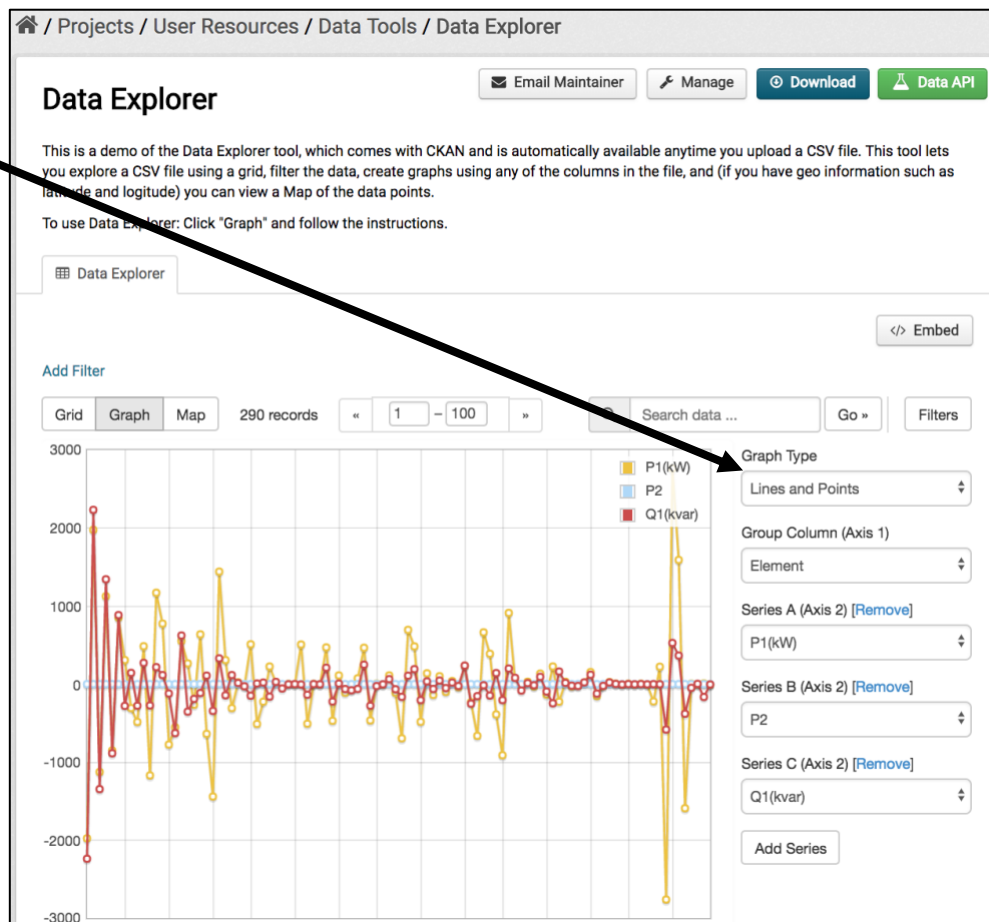
_id	Element	Terminal	P1(kW)	Q1(kvar)	P2	Q2	P0	Q0	P_Normal	Q_Normal
1	Line.INF...	1	-1976.7	-2238.1	0	0	0	0	0	0
2	Line.INF...	2	1976.8	2232.5	0	0	0	0		
3	Line.A1...	1	-1125.8	-1345.7	0	0	0	0	0	0
4	Line.A1...	2	1125.8	1342.9	0	0	0	0		
5	Line.A1...	1	-851	-886.8	0	0	0	0	0	0
6	Line.A1...	2	851	884	0	0	0	0		
7	Line.A1A...	1	307.6	-279.3	0	0	0	0	0	0
8	Line.A1A...	2	-307.6	143.6	0	0	0	0		
9	Line.A1A...	1	-485	-277.2	0	0	0	0	0	0
10	Line.A1A...	2	485	271.6	0	0	0	0		
11	Line.A1A...	1	-1170.1	-273.6	0	0	0	0	0	0
12	Line.A1A...	2	1170.2	217.1	0	0	0	0		
13	Line.A1A...	1	774.7	115.3	0	0	0	0	0	0
14	Line.A1A...	2	-774.7	-120.9	0	0	0	0		
15	Line.A1A...	1	-553	-628.1	0	0	0	0	0	0
16	Line.A1A...	2	553	622.5	0	0	0	0		
17	Line.A1B...	1	264.9	-353	0	0	0	0	0	0
18	Line.A1B...	2	-264.8	-189.9	0	0	0	0		
19	Line.A1B...	1	637.3	-114.2	0	0	0	0	0	0
20	Line.A1B...	2	-637.3	108.5	0	0	0	0		
21	Line.A1B...	1	-1440.8	-343.6	0	0	0	0	0	0
22	Line.A1B...	2	1440.9	329.5	0	0	0	0		

CSV Table (Grid): This is the default view of any CSV file. The headers from the file are also brought in. Each column is adjustable in size and can be filtered ascending or descending.

To change the display type for the CSV, click the buttons just above and to the left of the table.

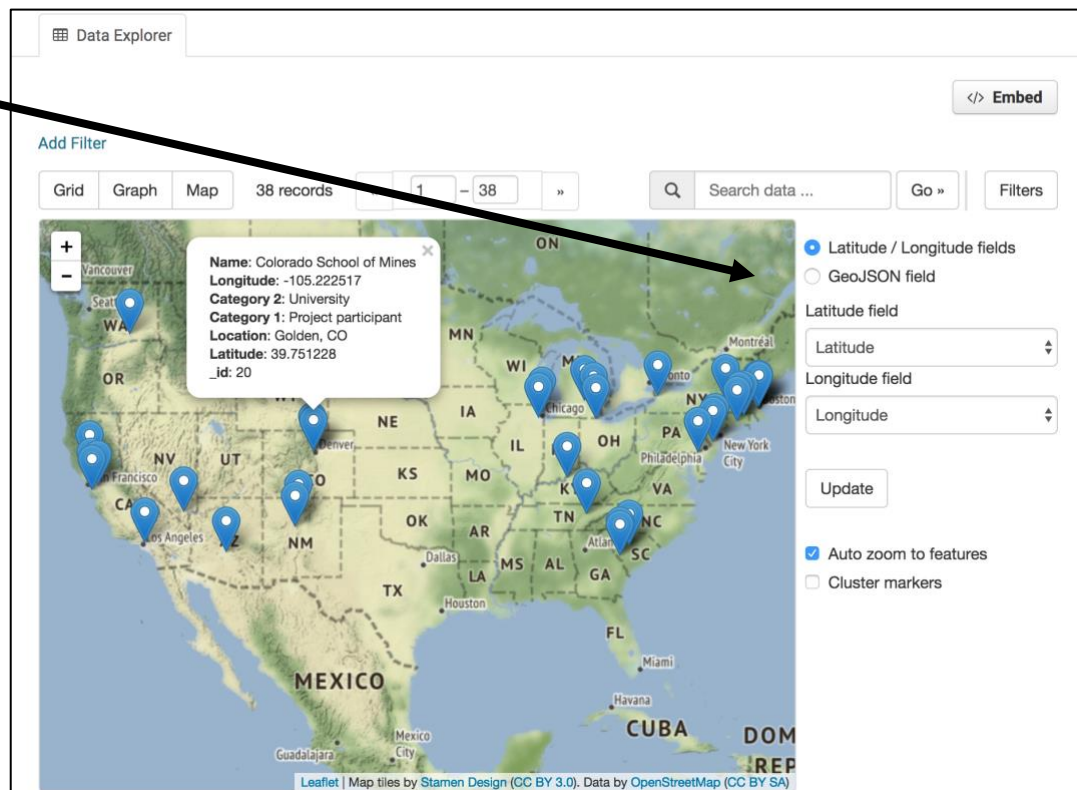


CSV Plot (Graph): The user needs to connect the columns in the table to the axis. This interface is on the right side in the Graph view.



CSV Map: The Data Explorer will automatically detect if the file has Latitude and Longitude header names and use those by default for the geographic coordinates.

Clicking any of the map markers will display all the data for the record associated with that geographic position.

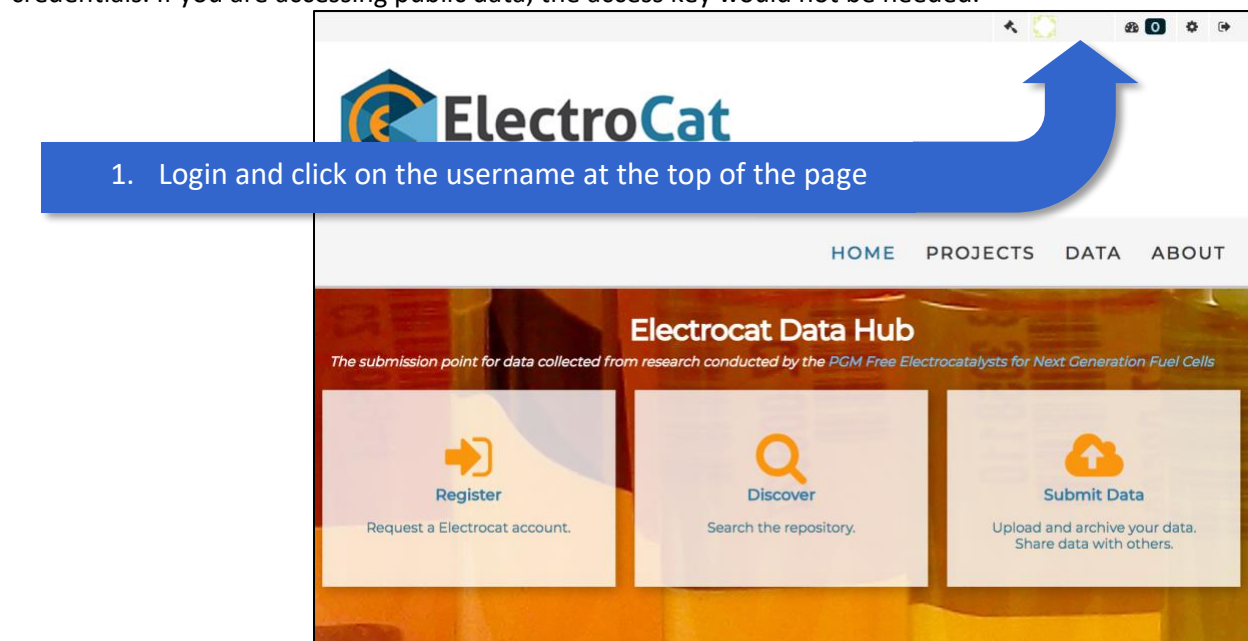


Accessing the Data through the API

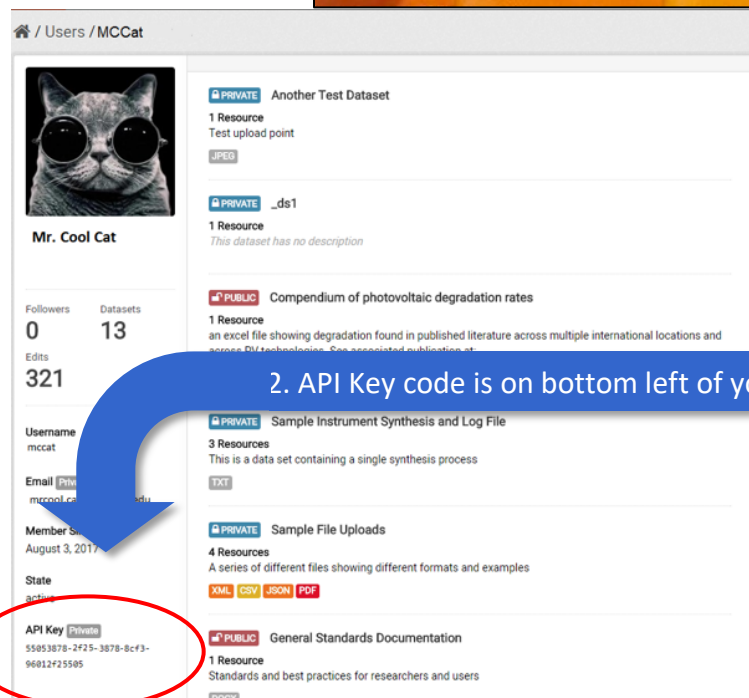
Data within the data hub can be retrieved with the Application Programmatic Interface (API) that is available as part of the data hub infrastructure. Secured, non-public data will still need to be accessed with credentials, but this can be passed as part of the API call. Using the API can be useful if you need to **access and download multiple files or datasets** on a regular basis in order to be processed through a software pipeline or tool set (e.g. Mathematica, Igor, Origin, etc.). The ElectroCat data hub uses the native CKAN API, which is built on a RESTful interface.

The first step to being able to access secure data through the API is to get your programmatic credentials. If you are accessing public data, the access key would not be needed.

1. Login and click on the username at the top of the page



2. API Key code is on bottom left of your User page. Write it down!



For further information on using the API and your API key for secure data access please see:
[CKAN API Documentation](#)

For Project Principle Investigators (PIs)

Adding new members to a project

Contact the [administrator](#) to request to add a new member to a project. If the new member has not registered, the PI may send an email invitation to the new member (Cc: emnadmin@nrel.gov) inviting them to register on the data hub.

There are two main levels of access:

1. **Member** – Read only access to data in the project
2. **Editor** – Read and add datasets or data to existing datasets

Changing permissions for a member of a project

Occasionally a project PI may request to adjust permissions for a member; giving or removing permissions to add or read data. Contact the [administrator](#) to adjust permissions for any member of a project.

Removing a project member

Contact the [administrator](#) to request to remove a user from a project.

Questions and Feedback?



- If you cannot log on to the website,
- Have problems with your data, and/or
- Have suggestions for improvement

Please contact the [administrator](#).