DuraMAT
Durable Module Materials Consortium

A Researcher's Guide to the DuraMAT Data Hub:

How To Make It Work for Your Project and Research

Robert R. White

Aug 10, 2020

Energy Materials Network
U.S. Department of Energy
NREL
NATIONAL RENEWABLE ENERGY LABORATORY
Sandia National Laboratories
BERKELEY LAB
SLAC NATIONAL ACCELERATOR LABORATORY

What is a Data Hub?

Simply,

A software platform that provides a virtual laboratory with a robust method of secure data management, data exploration, and dissemination, allowing the remote researchers to work efficiently and effectively.

Organizationally,

The Data Hub is part of Capability one: Central Data Resource.

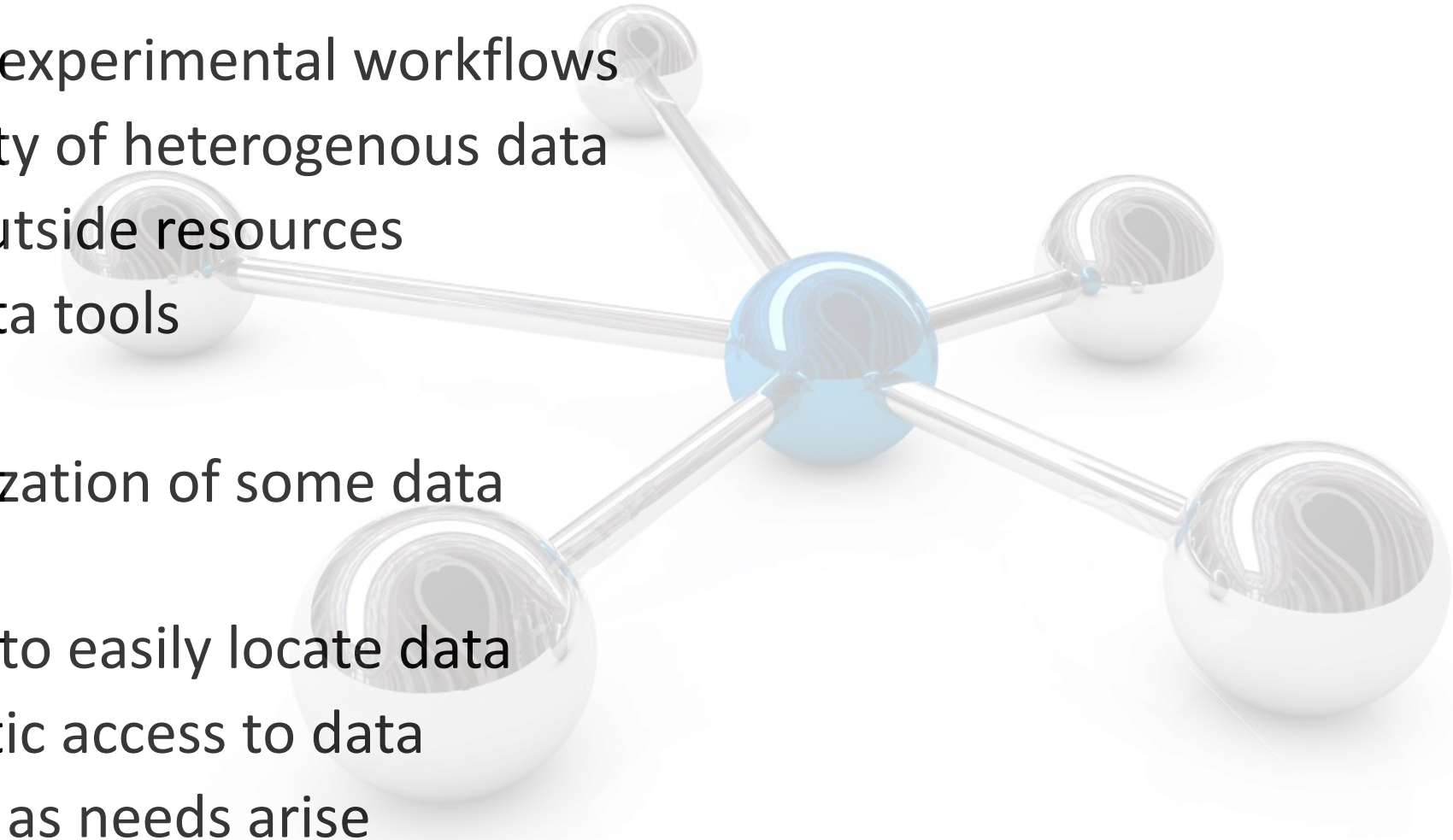
Technically,

The Data Hub is series of customizable software elements integrated to provide a data repository and website, deployed in the cloud.

Follow along or more help at: https://datahub.duramat.org/dataset/help-and-tutorial

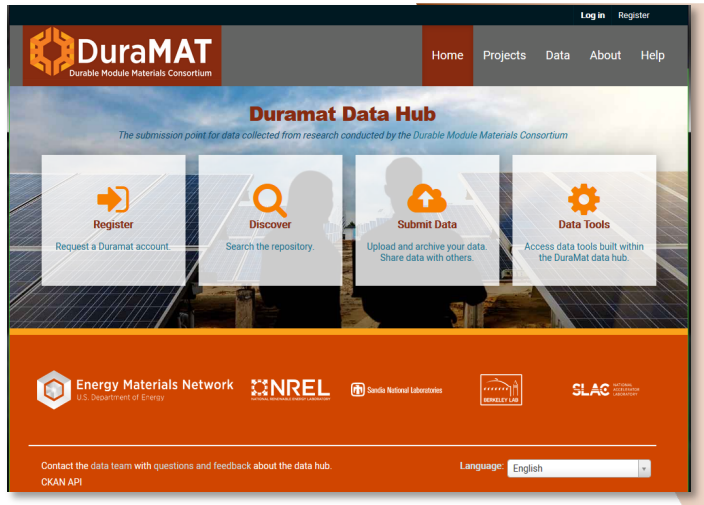
What Can the Data Hub Do?

- Support a variety of experimental workflows
- Archive a wide variety of heterogenous data
- Support linking to outside resources
- Archive or link to data tools
- Capture metadata
- Provide quick visualization of some data
- Secure web portal
- Queryable interface to easily locate data
- Provide programmatic access to data
- Allow customization as needs arise



Where is the Data Hub?

https://datahub.duramat.org



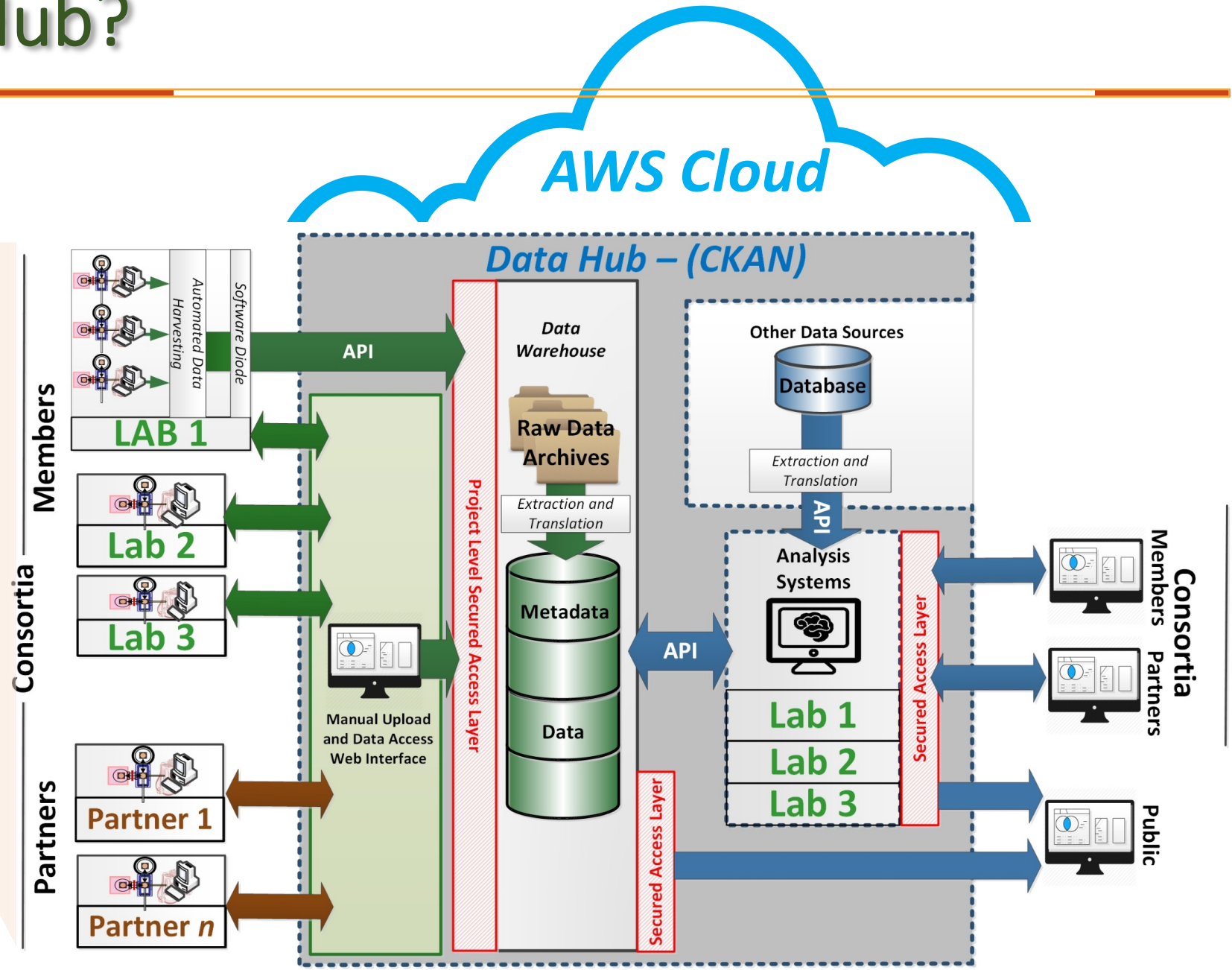
Ancillary sites

https://pvtools.lbl.gov

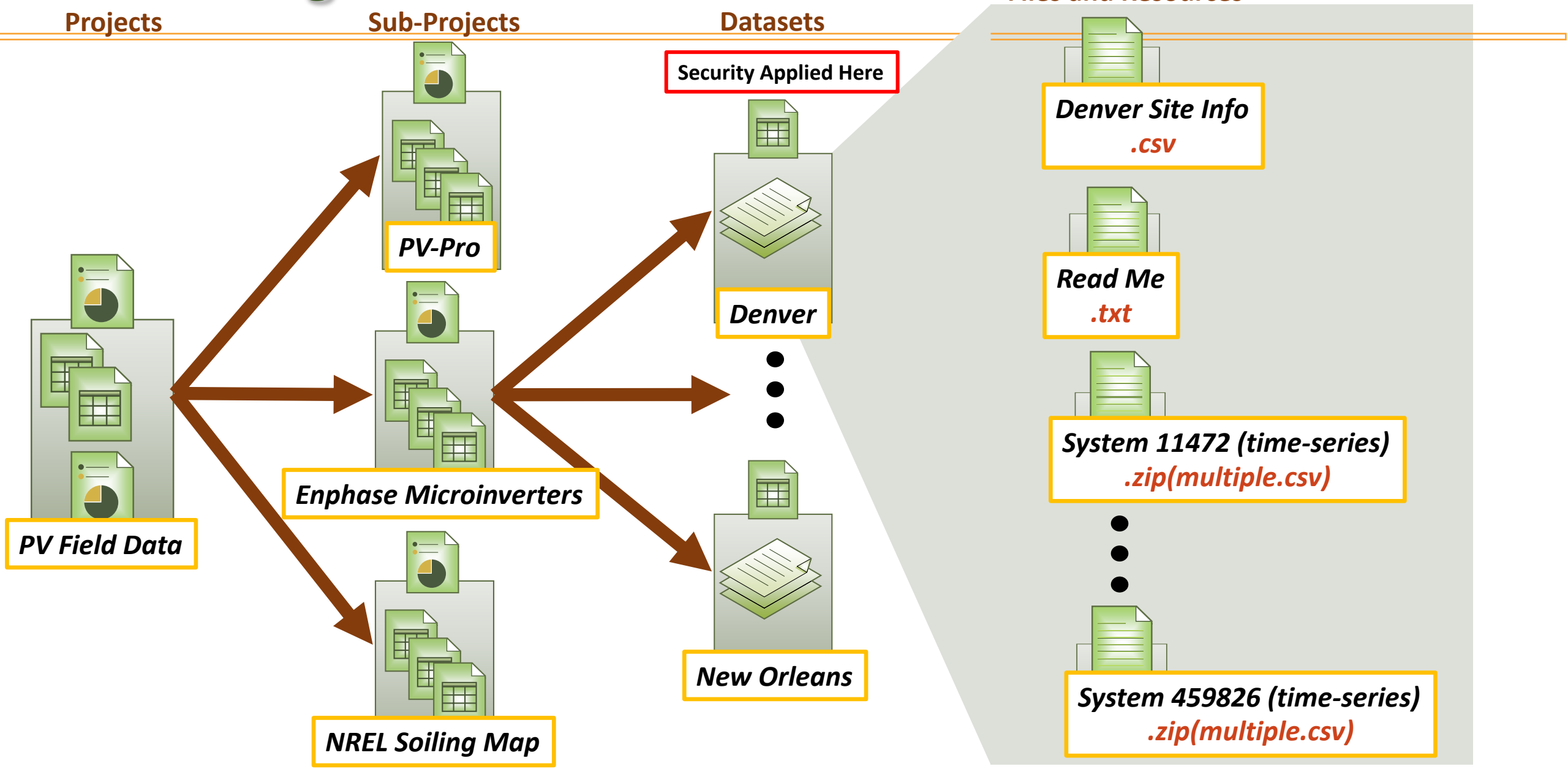
https://pvdata.duramat.org

https://github.com/DuraMAT

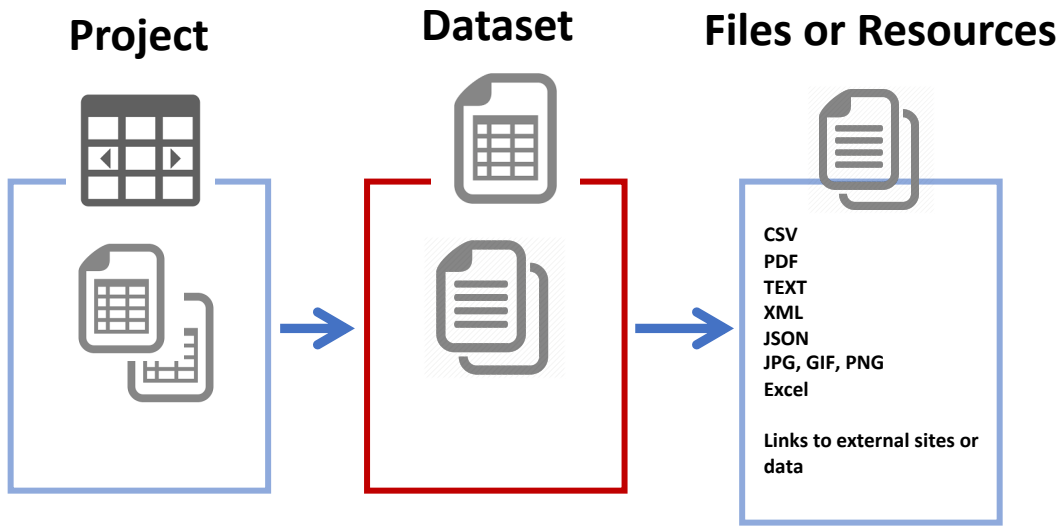
https://github.com/EMN-Data



Data Management Architecture



Data Security



***Public Access*:**
The public can see project names, descriptions and abstracts

Team Member Access:
Authenticated members can see all project information, and contained datasets and files

- Researcher logins via Cognito
  - Login access works for all datahubs
- Data security applied at dataset level
- **All** data uploads as "private"
- Three Project Member permissions
  - Member – Can see all data in project
  - Editor – Can add new datasets and data
  - Admin – Can add new members to project

- Data uploads only from:
  - DuraMAT researchers
  - Vetted data sources

  (no general public data)
- Public data release under administrative control, requires authorization.

What's in the Data Hub

For the public:

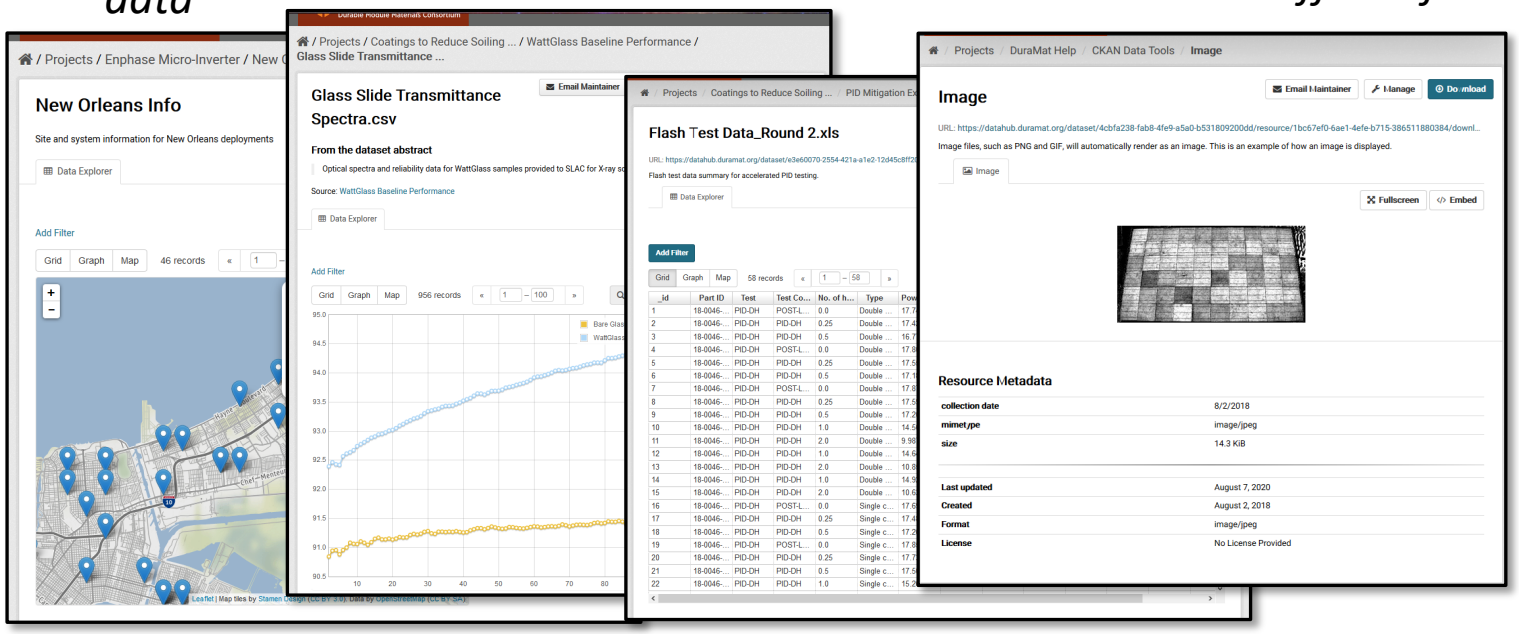
- All project descriptions and abstracts
- Data tools
- Publicly available datasets

1. *Albedo Data for Bifacial Systems*
2. *Coatings to Reduce Soiling and PID Losses*
3. *NREL Soiling map and supporting data*
4. *Bifacial Experimental Single-Axis Tracking Field data*

For project members:

- Project datasets
- Data tools
- All publicly available datasets

1. *Back side defect imaging in crystalline silicon photovoltaic modules*
2. *Identifying Degradation Mechanisms in Fielded Modules using Luminescence and Thermal Imaging*
3. *Combined Accelerated Stress Testing*
4. *Effect of Cell Cracks on Module Power Loss and Degradation*



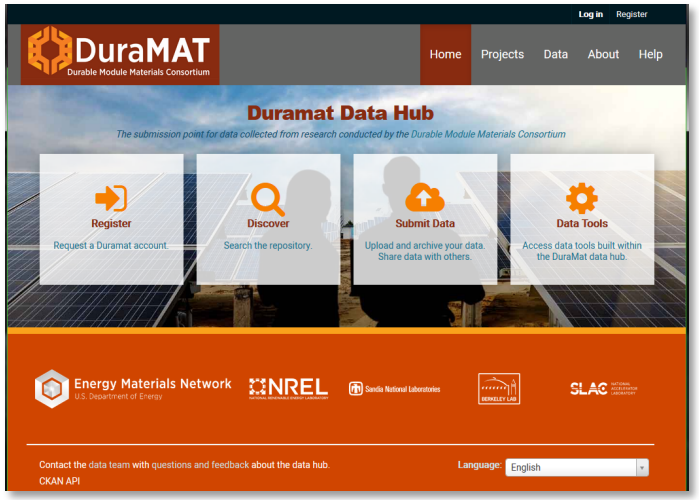
Data Tool Links:

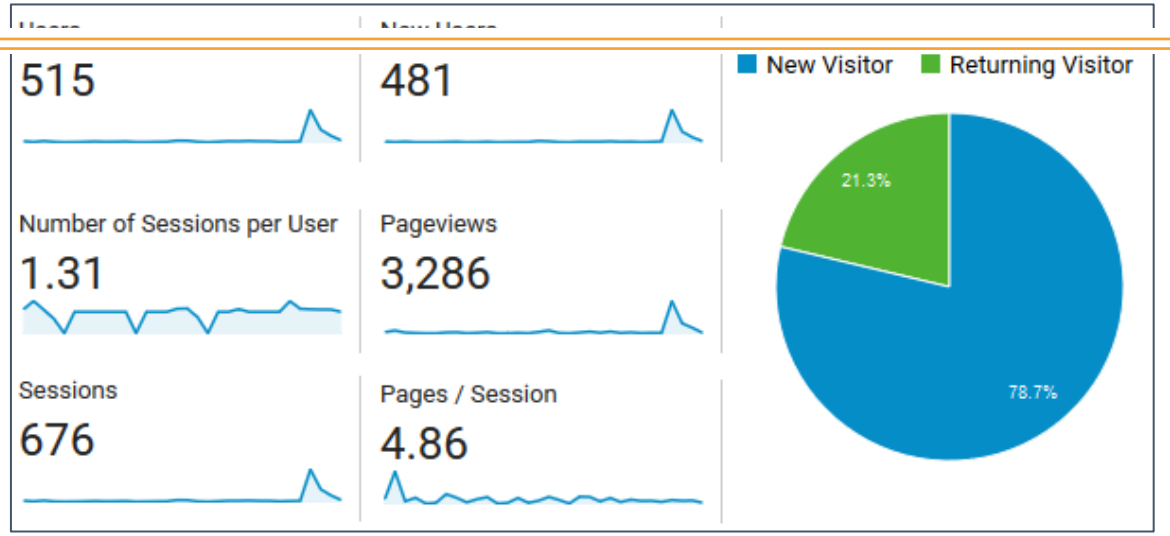
1. Rd Tools
2. Photovoltaic String Calculator
3. PV Climate Zone Stressors
4. IV Curve Corrections
5. Clear Sky Detection

….and more

Is Anybody Using It?

Google analytics July 1 –Aug 10, 2020:





Data Hub Stats

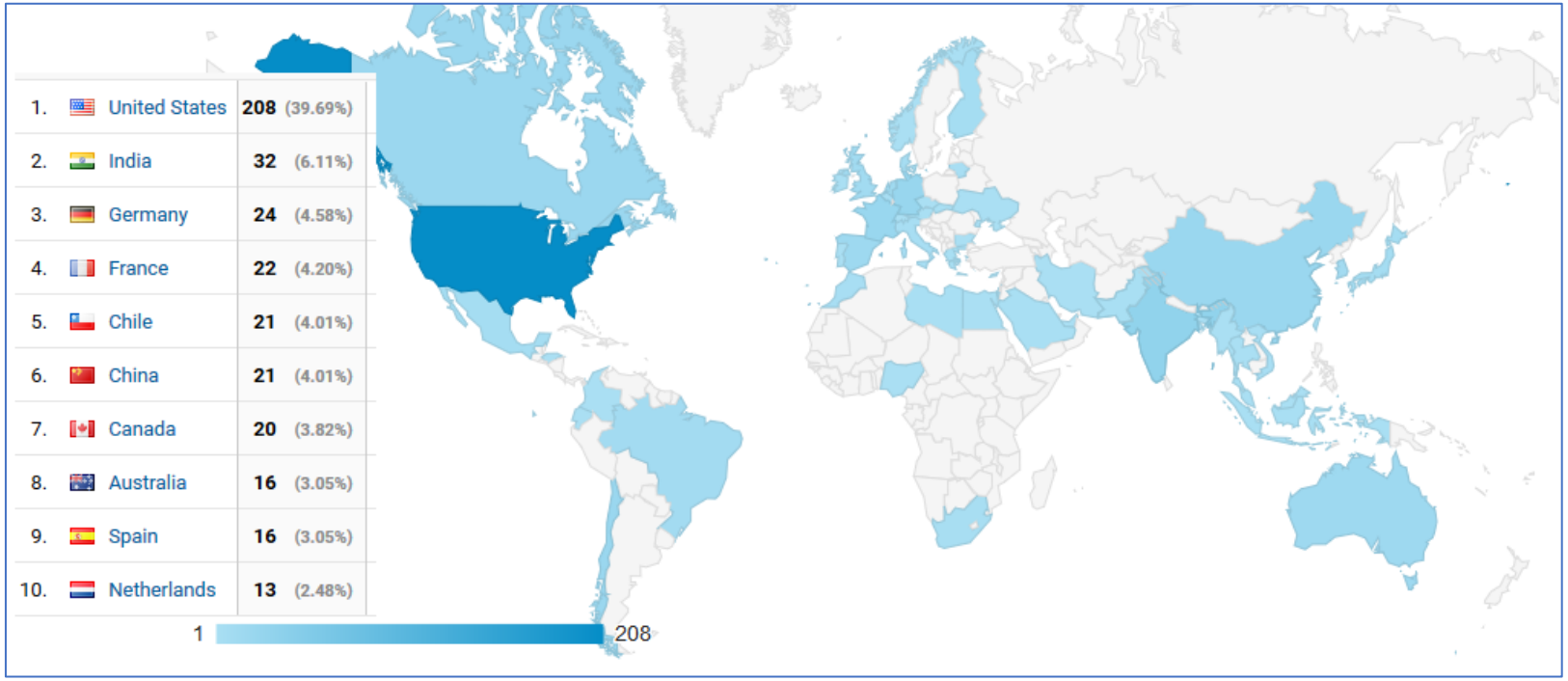
137 DuraMAT Users

63 Projects

128 Datasets

70 Public Datasets

2120 Files and Resources



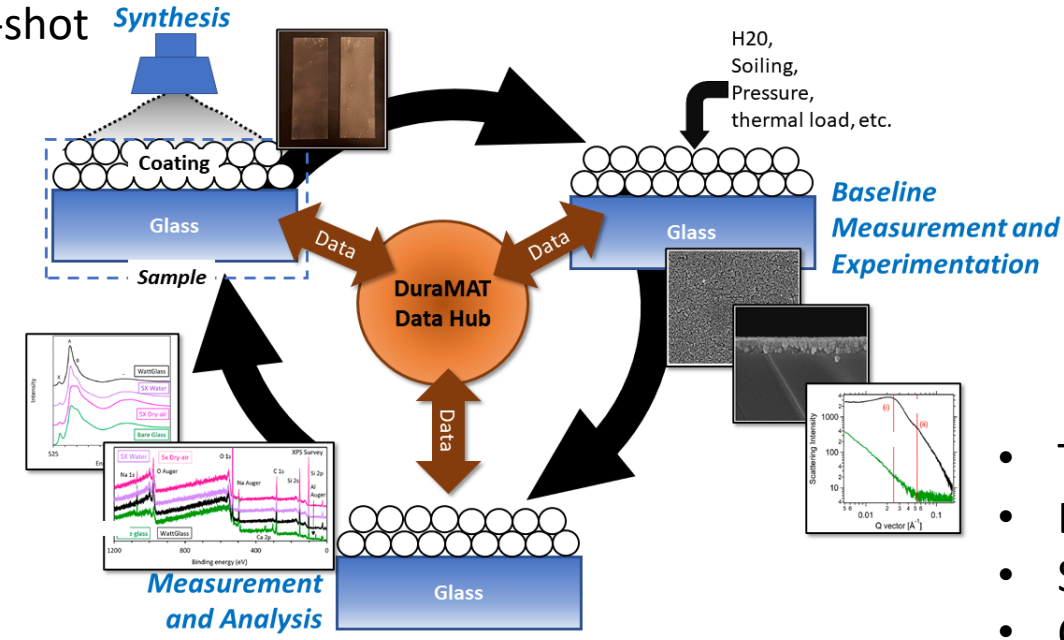
Making it work

How do I set up my project?

A. Design the Project Archive

Workflow

1. Round robin experiments
2. Running data archive
3. One-shot



Data

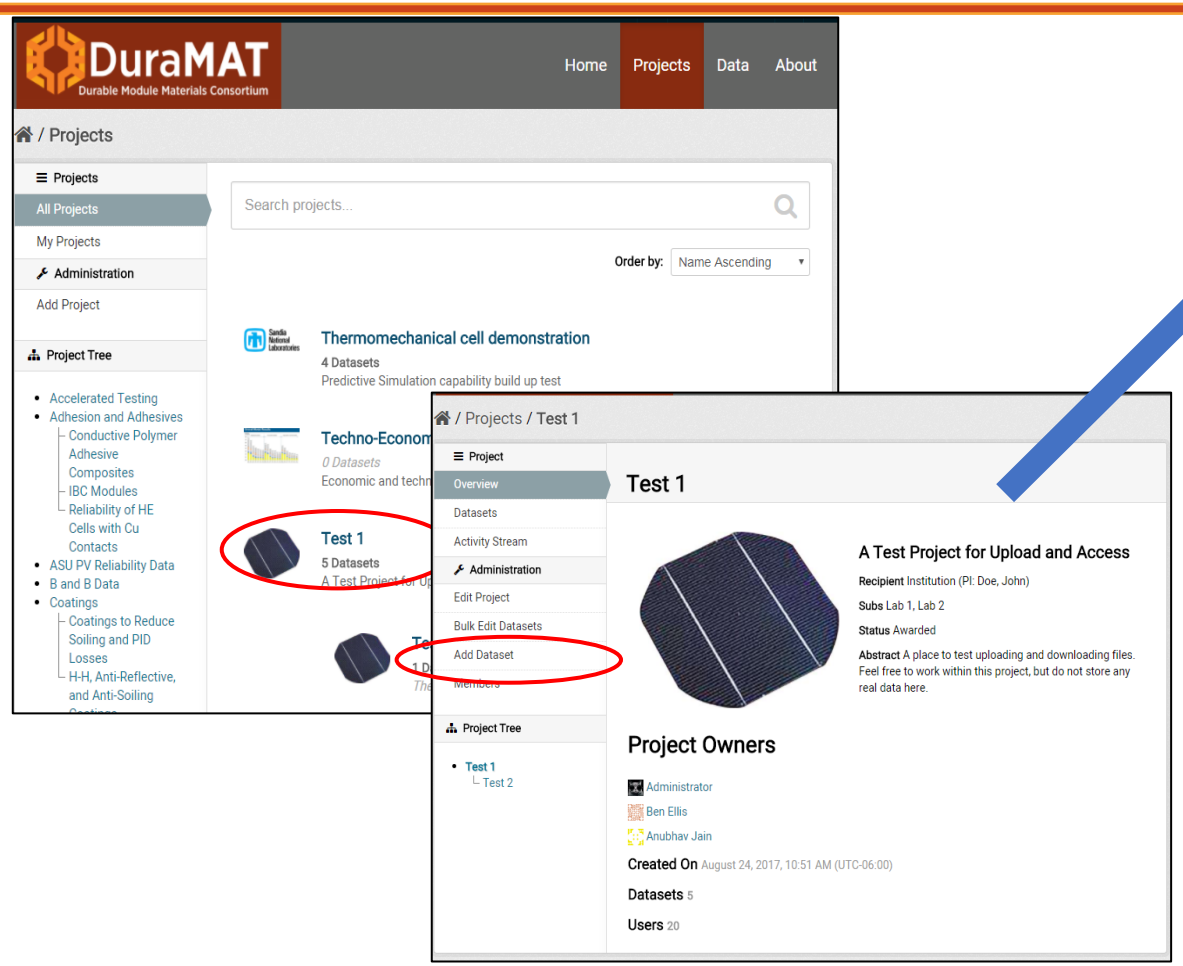
1. Instrumentation
2. **Metadata**
3. Validation / QA
4. Moving data
5. Other team access

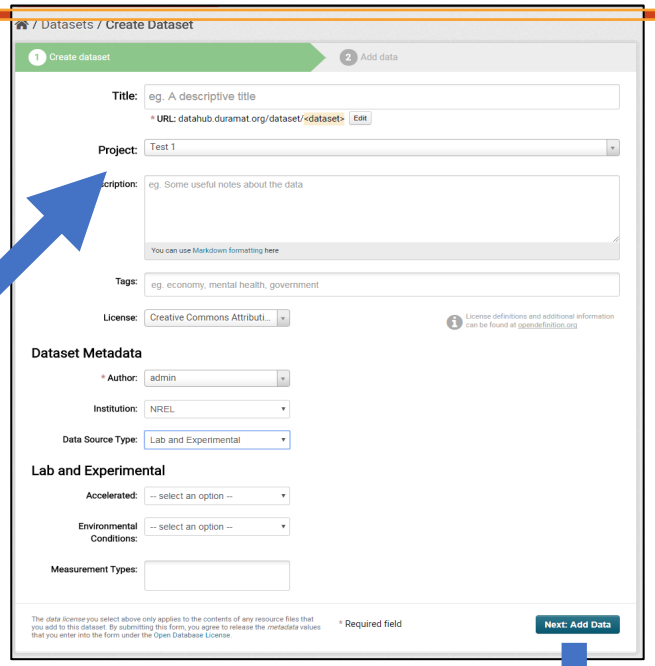
B. Project Description Form

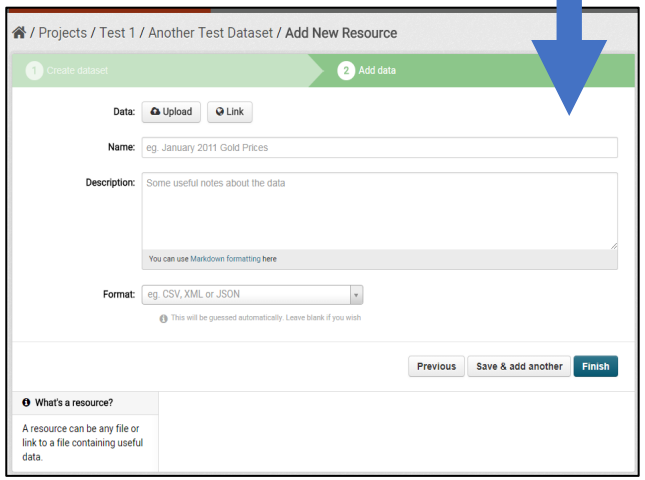
https://datahub.duramat.org/dataset/new-project-request

- Category of project
  - Accelerated testing
  - Adhesions and adhesives
  - Coatings
  - Encapsulants and barriers
  - Mechanics and cracking
  - Module forensics
  - Data analysis and quality assurance
  - Predictive simulation
  - Modeling
  - Field Data Studies
- Title
- Lead Organization and PI name
- Supporting organizations
- Current Status (Awarded, Complete, in progress)
- Abstract (Detailed description of the project and its goals)
- Access List (What DuraMAT members need to get to the data)
  - What kind of data access does each member need:
    - Member, Editor, or Admin

How do I put my project data in the Data Hub?







1. Select the project
2. Click “Add dataset”
3. Fill out interactive form
4. Click button “Add Data”
5. Fill out interactive form

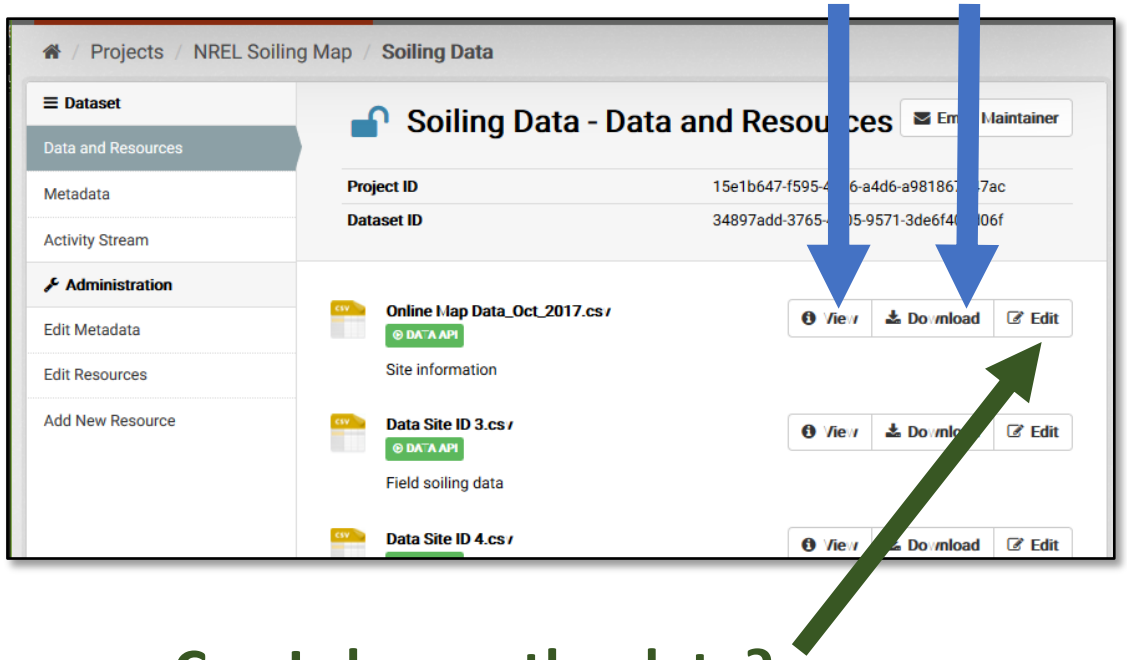
Or you can use the API and perform the same steps programmatically.

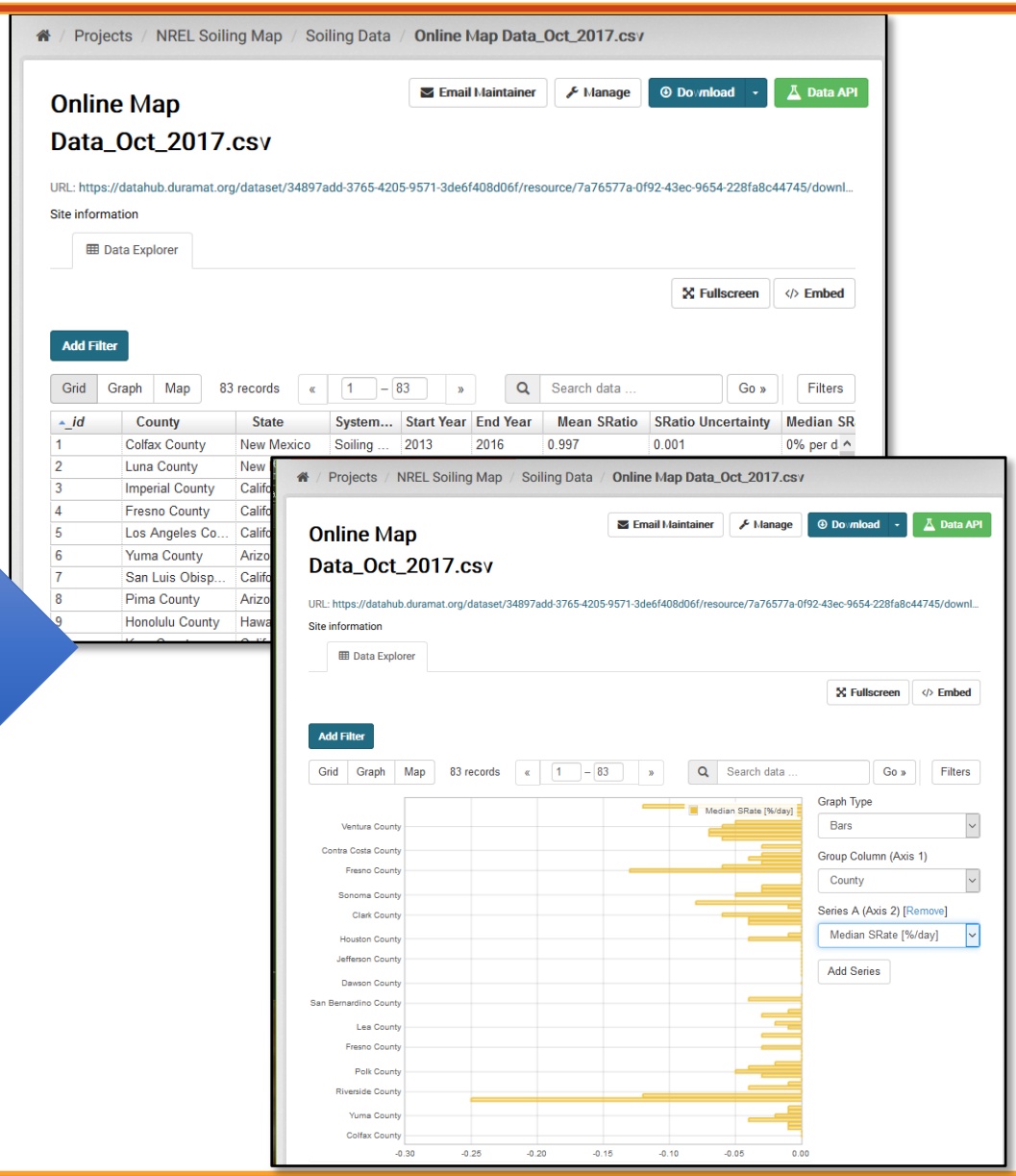
A new multiple file (bulk) upload feature is in test and will be available next month.

How do I get my project data?

Using the Website

1. Log into the Data Hub (For private project data only)
2. Click on Projects and select the project from the list
3. On left hand menu, click “Datasets”
4. Click the dataset you want to access from the list
5. Select the data you wish to see (“View”) or download





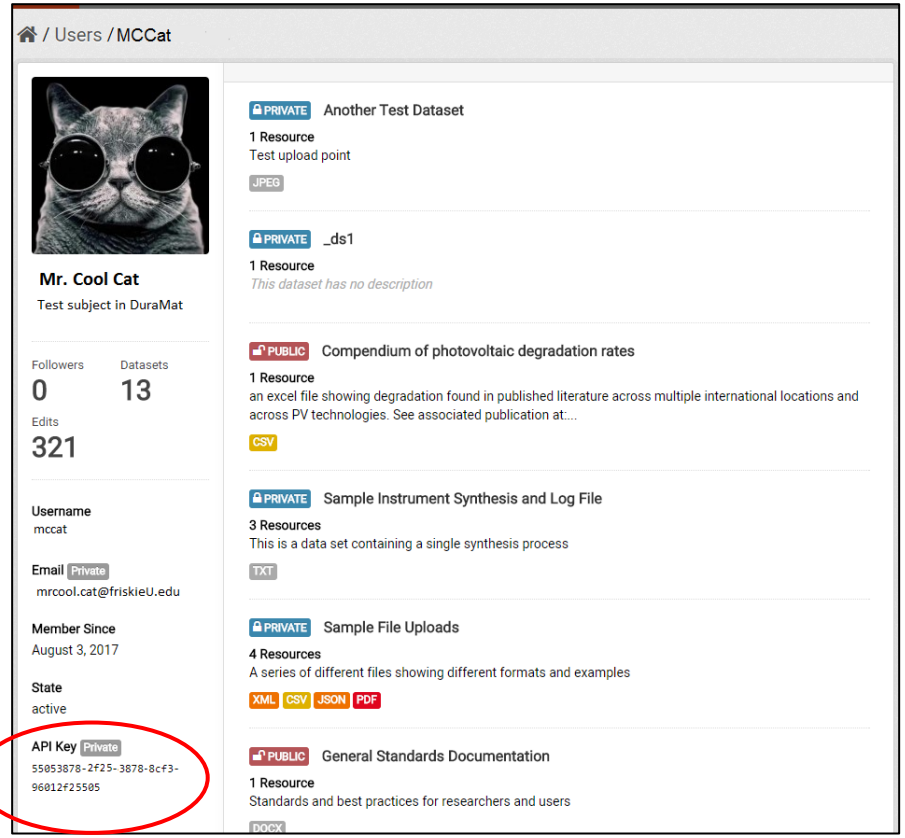
Can I change the data?

Click “Edit” next to the name of the file

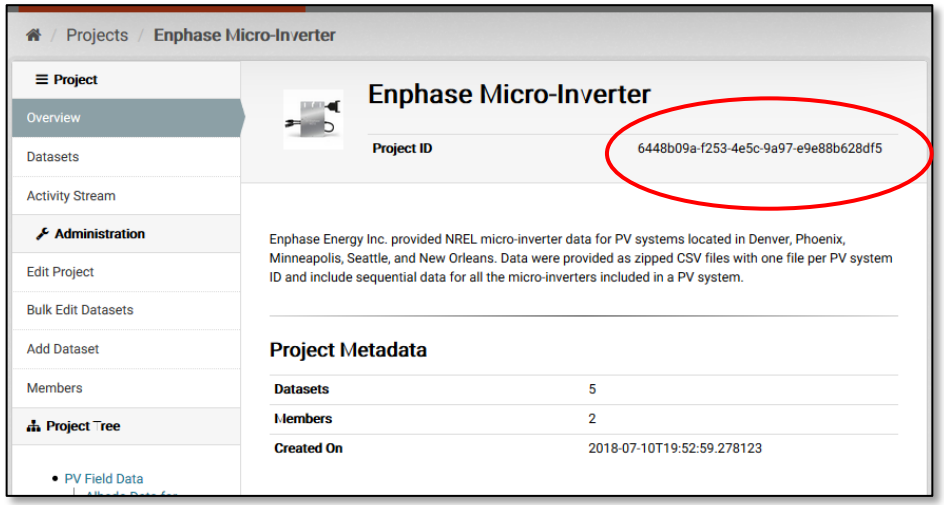
Accessing data without the Web App

All items stored within the Data Hub is assigned a Universally Unique Identifier (UUID); projects, datasets, data files... ***and even registered users***

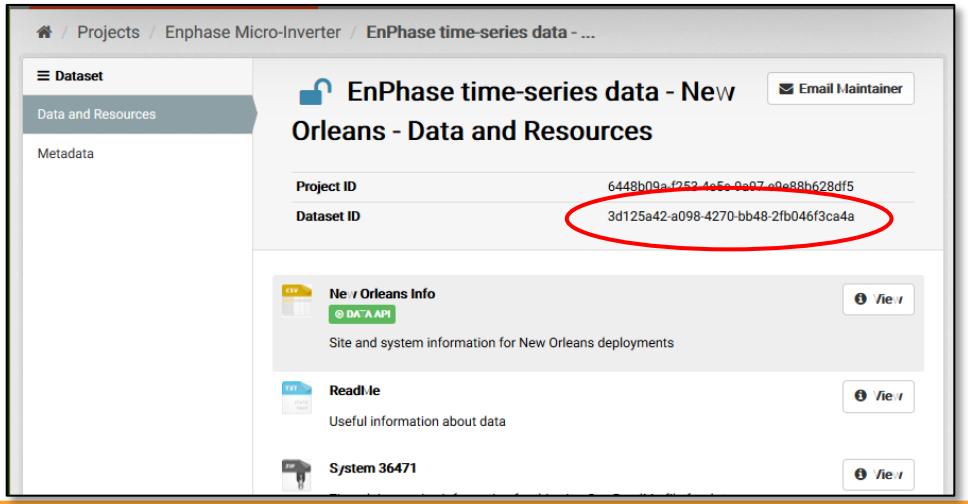
user UUID (a.k.a API Key) – User Page



project UUID – Project overview page
Project datasets page

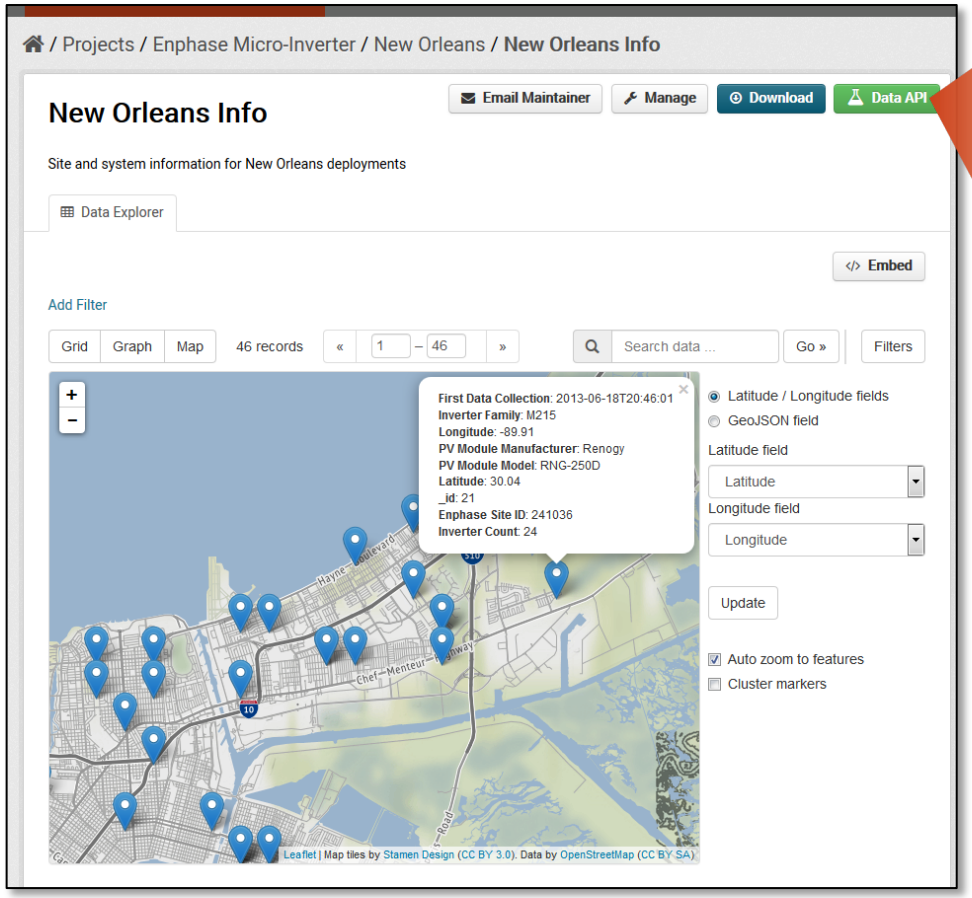


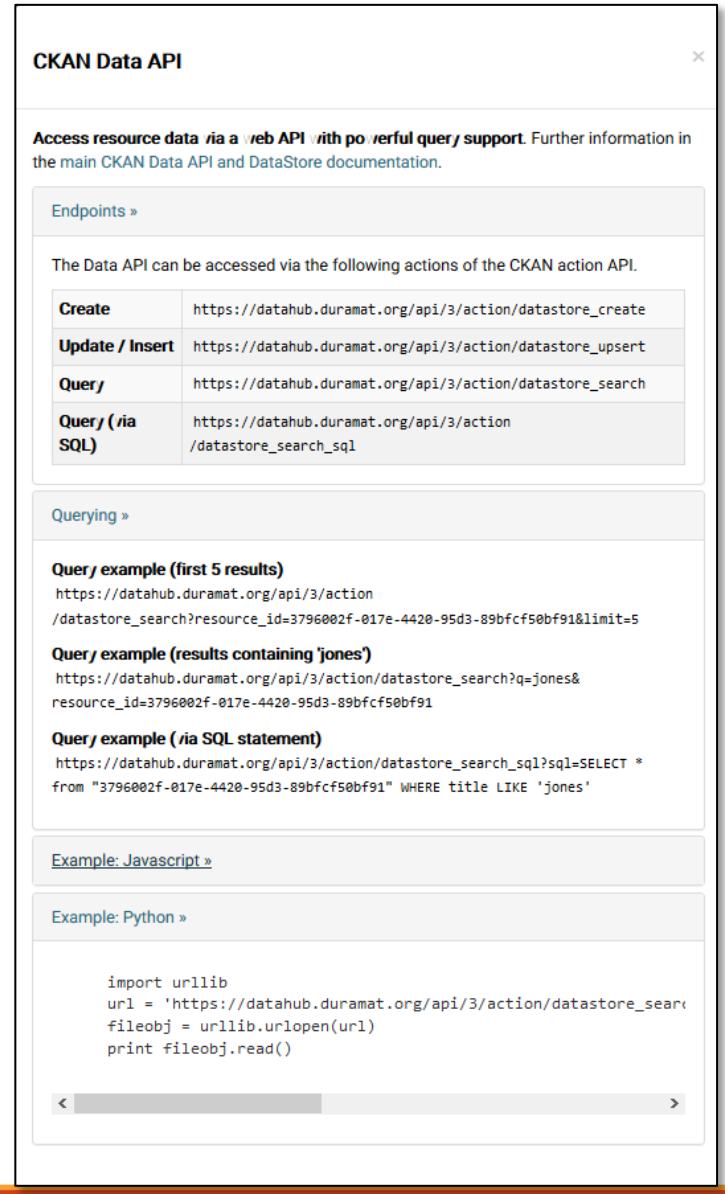
dataset UUID – Dataset page



Accessing the data file through API

All files are accessible for download from the API, but CSV files are special, and the API will allow for querying and accessing records in the CSV file.





API Demo

Need more help?

Go to the Data Hub https://datahub.duramat.org

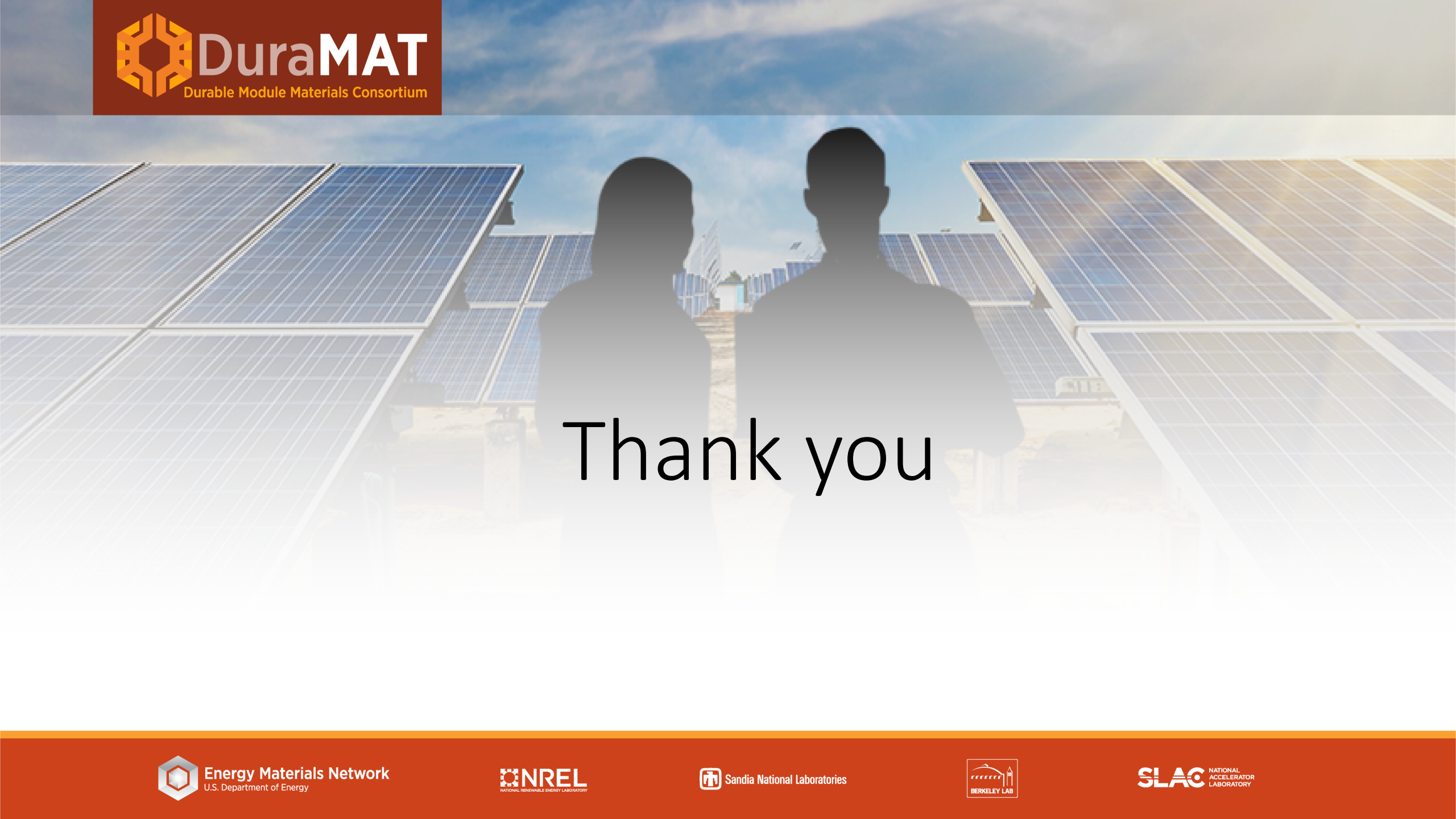
Go to Projects > DuraMAT Help > Datasets > Help and Tutorial

- Data Hub User's Manual – Covers all of this and more
- DuraMAT Data Hub Api Walkthrough - Another copy of the API iPython notebook
- Video Tutorial

Acknowledgements:

(the developers, supporters and early adopters)

Kris Munch
Courtney Pailing
Chris Webber
David Ragar
David Whiteside
Teresa Barnes
Nick Wunder
Nalinrat Guba
Michael Bahl
Jainli Gu
Laura Schelhas
Stephanie Moffitt
Anubahv Jain
Todd Karin
Ben Ellis
Birk Jones
Cameron Stark
Margaret Gordon


DuraMAT
Durable Module Materials Consortium
Thank you
Energy Materials Network
U.S. Department of Energy
NREL
NATIONAL RENEWABLE ENERGY LABORATORY
Sandia National Laboratories
BERKELEY LAB
SLAC
NATIONAL ACCELERATOR LABORATORY