

### 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











### 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?

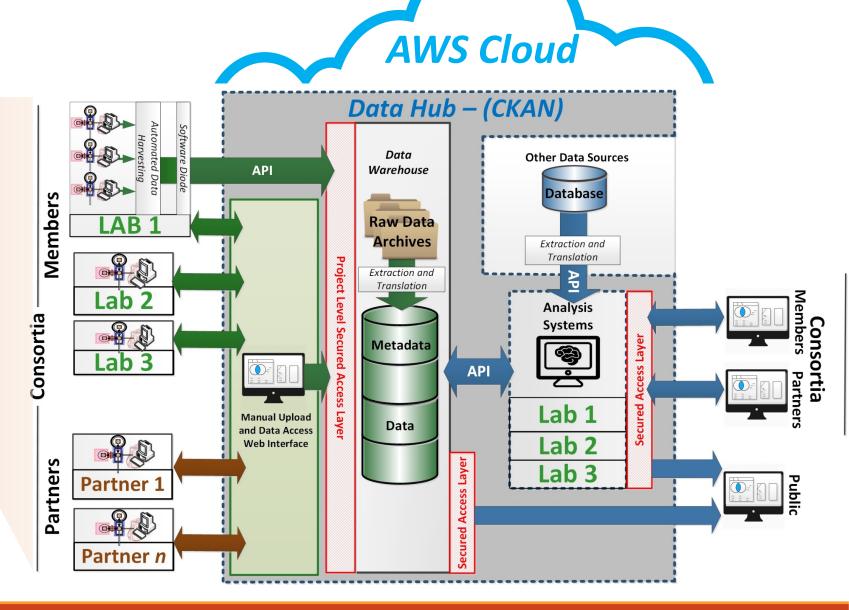
#### https://datahub.duramat.org

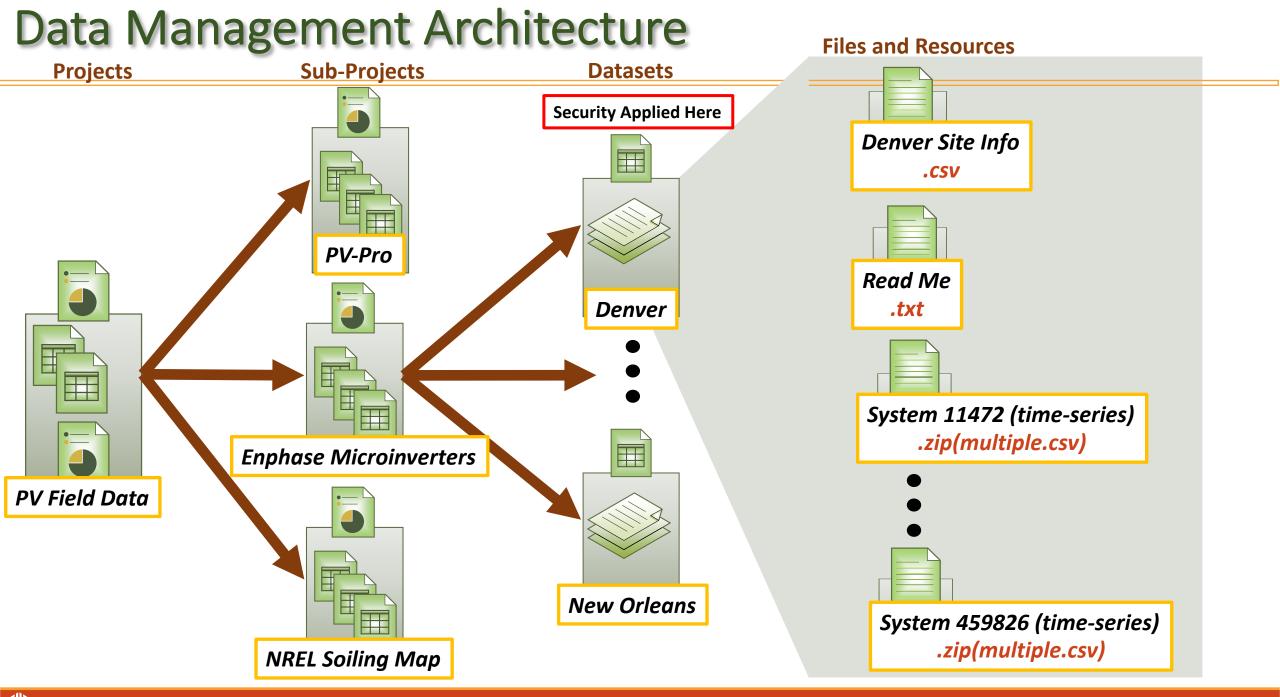


#### Ancillary sites

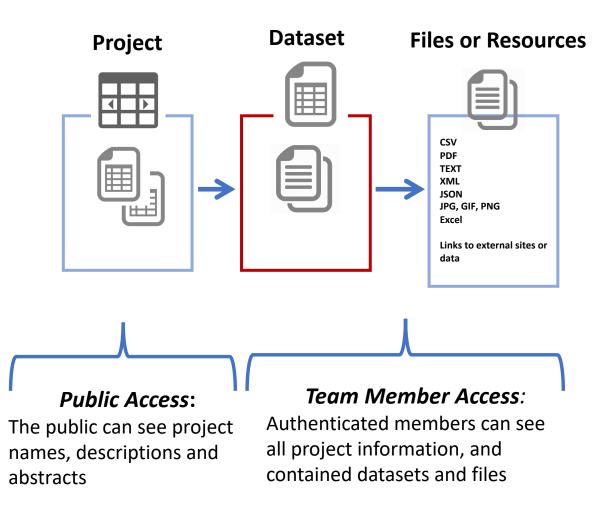
https://pvtools.lbl.gov https://pvdata.duramat.org https://github.com/DuraMAT

https://github.com/EMN-Data





### **Data Security**



- 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

🛠 / Projects / Enphase Micro-Inverter / New (	A / Projects / Coatings to Reduce Soiling / WattGlass Baseline Glass Slide Transmittance	Performance /	🕷 / Projects / DuraMat Help / CKAN Data Tools / Image	
New Orleans Info Site and system information for New Orleans deployments Data Explorer	Glass Slide Transmittance     Spectra.csv      From the dataset abstract     Optical accert and reliability data for WattGlass samples provided to SLAC for X-ray or Source: WattGlass Baseline Performance	Projects Coatings to Reduce Solling PID Mitigation E     Flash Test Data_Round 2.xls     With hype/Mathub duramet org/bitsast/viele/0070 2554-271e artic/12/04508175     Plash test data summary for accelerated PD testing.     Exact Laboratory	URL: https://distahub.duramat.org/distaset/4cbfa238-fab8-4fe9-a5a0-b331809200dd/resource/1bc57ef0-6ae1-4efe-b715-386511 Image Eles, such as PNG and GIF, will automatically render as an image. This is an example of how an image is displayed.	Do mload
Add Filter Grid Graph Map 46 records < 1 - + -	ED bata Explorer           Add Fiter           Grid         Gragen           8.0           8.0           8.0           8.0           8.0           8.0           8.0           8.0           8.0           8.0           8.0           8.0	Add Tener         Image: Construct on the second secon	167 178 175 175 175 175	
	935 939 925 920 955 910 925 910 920 920 920 920 920 920 920 92	B         H 8005E         PDDH         PDDH         0.25         Double         17           9         H 8005E         PDDH         PDDH         0.25         Double         17           9         H 8005E         PDDH         PDDH         PDDH         16         Double         11           10         H 8005E         PDDH         PDDH         PDDH         20         Double         14           11         H 8005E         PDDH         PDDH         PDDH         20         Double         14           13         H 8005E         PDDH         PDDH         PDDH         20         Double         14           15         H 8005E         PDDH         PDDH         PDDH         20         Double         14           16         H 8005E         PDDH         PDDH         PDDH         20         Double         14           16         H 8005E         PDDH         PDDH         PDDH         20         Double         16           17         H 8005E         PDDH         PDDH         20         Double         16         90         16         90         90         90         90         90         90         90	175         collection date         8/2/2018           186         mimetype         image/jpeg           186         size         143 KB           186         L43 KB         143 KB           186         Size         143 KB           187         Forested         August 7, 2020           188         August 7, 2020         143 KB           189         Kested         August 7, 2020           189         Lessee         Magest 2, 2018	
C Leafet   Map ties by Stamen C	10 20 30 40 50 60 70 80	<ul> <li>es periorio numeri n Numeri numeri numer Numeri numeri nume</li></ul>		

### **Data Tool Links:**

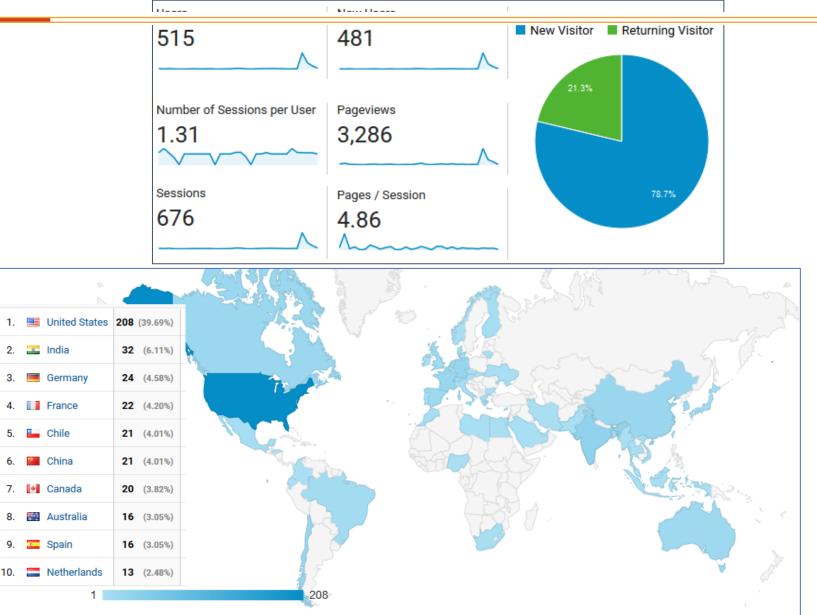
- 1. Rd Tools
- 2. Photovoltaic String Calculator
- 3. PV Climate Zone Stressors
- 4. IV Curve Corrections
- 5. Clear Sky Detection ....and more

## Is Anybody Using It?

*Google analytics* July 1 – Aug 10, 2020:

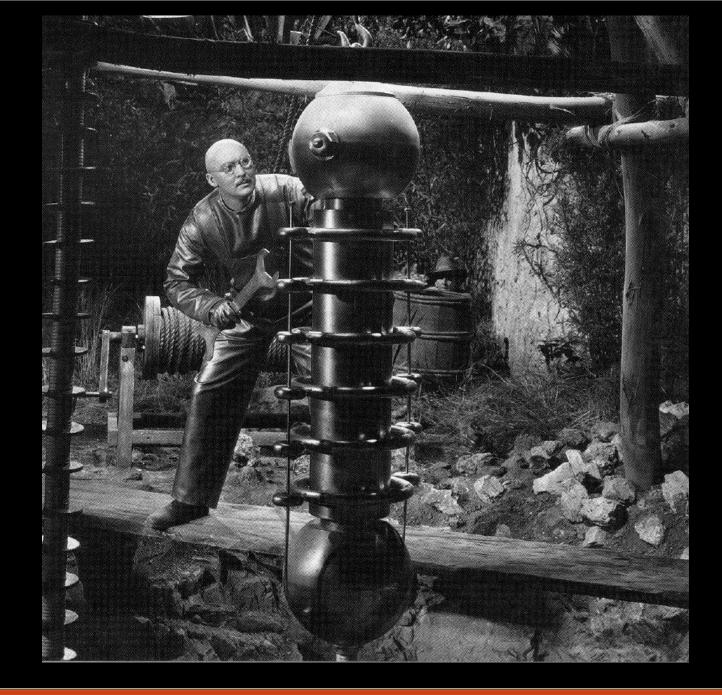


**Data Hub Stats 137** DuraMAT Users **63** Projects **128** Datasets **70** Public Datasets **2120** Files and Resources



2.

9.



## Making it work

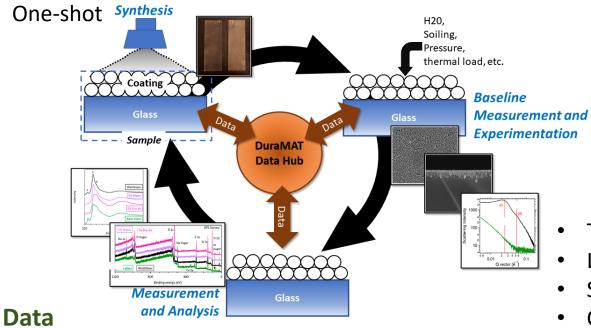
## How do I set up my project?

### A. Design the Project Archive

#### Workflow

- 1. Round robin experiments
- 2. Running data archive

3.



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

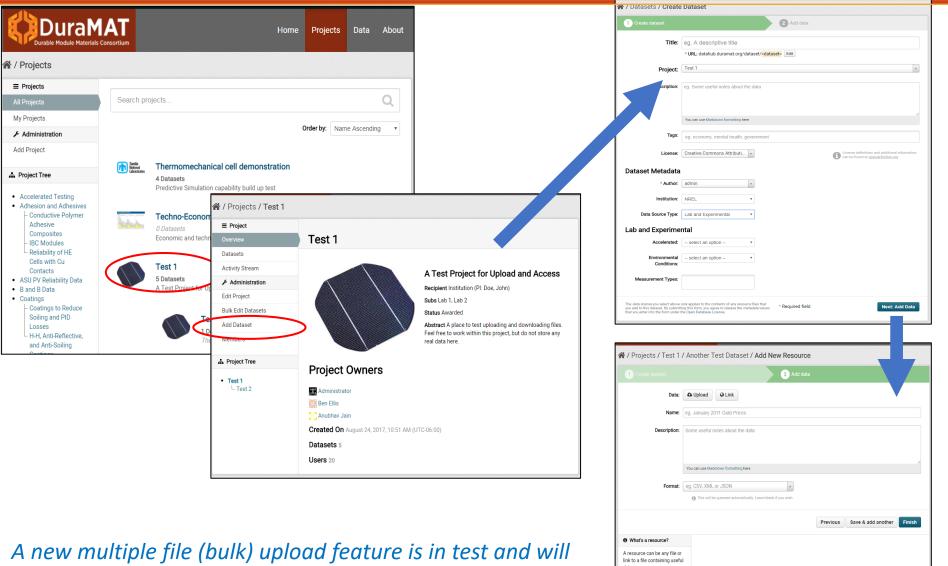
### **B. Project Description Form**

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)
  - $\circ~$  What kind of data access does each member need:
    - o Member, Editor, or Admin

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



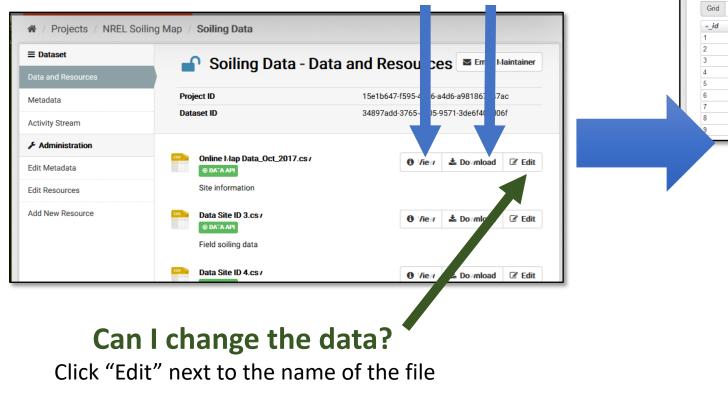
- 1. Select the project
- 2. Click "Add dataset"
- 3. Fill out interactive form
- 4. Click button "Add Data"
- 5. Fill out interactive form

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

## How do I get my project data?

#### Using the Website

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



	ojects / NREL Soil															
	ne Map a_Oct_2017	.csv				Email Main	tainer	<b>₽</b> Ma	inage	⊙ D	ov <b>nlo</b>	ad 🗸	👗 Da	ita API		
IRL: http	ps://datahub.duramat.o	rg/dataset/	34897add-3	3765-420	05-9571	-3de6f408d	06f/res	ource/7a76	6577a-01	f92-43e	c-9654	-228fa8c4	14745/do	ownl		
ite info	ormation															
I	Data Explorer															
										:	X Full	screen	Em	bed		
Add F	ilter														L	
Grid	Graph Map 8	33 records	« [·	1 –	83	»	Q	Search d	lata			Go »	Fil	ters	L	
_id	County	State	Sy	/stem	Start	Year End	Year	Mean S	SRatio	SRati	io Uno	ertainty	Media	an SR		
1	Colfax County	New Me	xico So	biling	2013	2016	6	0.997		0.001			0% pe	rd ^		
2	Luna County	New	A / Pro	jects /		Soiling Map	Soi	ling Data	Onlin		Data	Oct_2017	1.001			
3	Imperial County	Califo	mr / Plu	jects /	INREL C	sonnig wap	301	ing Data	Unini	emap	Data_	000_2017				
Ļ	Fresno County	Califo								nail Main	toiner	📕 Man		Dovnloa		A Data A
	Los Angeles Co.	Califo	Onli	ne M	ар						lamer	/ Mail	aye	9 DO MILOS	u .	A Data A
	Yuma County	Arizo														
		-	Data	a Oct	201	17 csv										
	San Luis Obisp	. Califo	Data	a_Oct	_ <b>20</b> 1	17.csv										
	Pima County	. Califo Arizo				<b>17.CSV</b> at.org/dataset	/34897a	ıdd-3765-420	5-9571-3d	le6f408d	06f/res	ource/7a765	i77a-0f92-4	43ec-9654	·228fa8c4	4745/downl
3		. Califo		ps://datahi			/34897a	ıdd-3765-420	5-9571-3d	le6f408d	06f/res	ource/7a765	i77a-0f92-4	43ec-9654	-228fa8c4	4745/downl.
3	Pima County	. Califo Arizo	URL: http Site infor	ps://datahi	ub.duram		:/34897a	ıdd-3765-420	5-9571-3d	le6f408d	06f/res	ource/7a765	i77a-0f92-≀	43ec-9654	-228fa8c4	4745/downl.
3	Pima County	. Califo Arizo	URL: http Site infor	ps://datahi rmation	ub.duram		:/34897a	ıdd-3765-420	5-9571-3d	le6f408d	06f/res	burce/7a765	77a-0f92-4			
7	Pima County	. Califo Arizo	URL: http Site infor	ps://datahi rmation	ub.duram		:/34897a	ıdd-3765-420	5-9571-3d	le6f408di	06f/res	ource/7a765	i77a-0f92-∙	43ec-9654		4745/downl.
3	Pima County	. Califo Arizo	URL: http Site infor	ps://datahu rmation I Data Exp	ub.duram		:/34897a	ıdd-3765-420:	5-9571-3d	le6f408di	06f/rese	ource/7a765	i77a-0f92			
	Pima County	. Califo Arizo	URL: http Site infor	ps://datahu rmation I Data Exp	ub.duram		:/34897a	ıdd-3765-420	5-9571-3d	Je6f408di	06f/res	ource/7a765	:77a-0f92			
	Pima County	. Califo Arizo	URL: http Site infor	ps://datahu rmation I Data Exp	ub.duram			idd-3765-420		Je6f408d	Q	Search da				
	Pima County	. Califo Arizo	URL: http Site infor	ps://datahu rmation I Data Exp	ub.duram lorer	at.org/dataset				) >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	٩	Search da	ta	S Fulk	Go »	♦ Embed
	Pima County	. Califo Arizo	URL: http Site infor Made Fi Grid	ilter	lorer Map	at.org/dataset				»	٩		ta	Sraph Type	Go »	Filters
	Pima County	. Califo Arizo	URL: http: Site infor Mdd Fi Grid	ilter	lorer Map	at.org/dataset				»	٩	Search da	ta	Sraph Type Bars	Go »	
	Pima County	. Califo Arizo	URL http Site infor Mdd Fi Grid	ilter	lorer Map	at.org/dataset				»	٩	Search da	ta	Graph Type Bars Group Colu	Go »	Embed Filters 1)
3	Pima County	. Califo Arizo	URL: http: Site infor Mdd Fri Grid Ve Contra	ilter Graph Costa Coun Fresno Coun	lorer Map	at.org/dataset				»	٩	Search da	ta	Sraph Type Bars	Go »	
	Pima County	. Califo Arizo	URL: http: Site infor Mdd Fri Grid Ve Contra	ilter Graph Countries Coun	lorer Map	at.org/dataset				»	٩	Search da	ta dayj [ C	Graph Type Bars Group Colu	Go » mn (Axis	♦ Embed Filters 1)
	Pima County	. Califo Arizo	URL: http: Site infor Mdd Fi Grid Vi Contra F So	ilter Graph Costa Coun Costa Coun Clark Coun	lorer Map ty ty ty ty	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A	Go » mn (Axis	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Mdf FT Grid Contra s So	ilter Graph Countries Coun	Ilorer Map ty	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A	Go » mn (Axis	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Mdd FT Grid Contra F So Hd Jef	ilter Graph Costa Coun Costa Coun Fresno Coun Clark Coun ouston Coun	Map Map ty	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Million Grid Contra 5 So Hit Jef	ilter Graph Graph Costa Coun Costa Coun Clark Coun ouston Coun	Map Iorer Map ty	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Million Grid Contra 5 So Hit Jef	Intermediate and a second seco	Map           Iv	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Million Grid Vi Contra 50 Son Ben San Benn	iter     inter     Counter     Conter	ub.duram           lorer           Map           ty	at.org/dataset				» »	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Million Grid Vi Contra 50 Son Hit Jet Di San Benni	Inter In	ub.duram           lorer           Map           ty	at.org/dataset				3 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Add Fi Grid Va Contra 5 o Son San Benn 5	Inter ps://datahi mation di Data Exp diter Graph Inter Communication Com	Map           fy	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Million Grid Via Contra F So Contra F So So Son Benni F Rive	Iter Graph Iter Graph Iter Company Com	Map           fy	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]
	Pima County	. Califo Arizo	URL: http: Site infor Grid Grid Vi Contra 5 5 5 5 5 5 5 5 5 5 5 7 7 7 7 7 7 7 7	Inter ps://datahi mation di Data Exp diter Graph Inter Communication Com	Map           hy           hy           y <td>at.org/dataset</td> <td></td> <td></td> <td></td> <td>»</td> <td>٩</td> <td>Search da</td> <td>ta dayj [ C</td> <td>Sraph Type Bars Group Colu County Series A (A Median S</td> <td>Go » mn (Axis xxis 2) [Re SRate [%/</td> <td>Filters 1) move]</td>	at.org/dataset				»	٩	Search da	ta dayj [ C	Sraph Type Bars Group Colu County Series A (A Median S	Go » mn (Axis xxis 2) [Re SRate [%/	Filters 1) move]

A / Proje

## Accessing data without the Web App

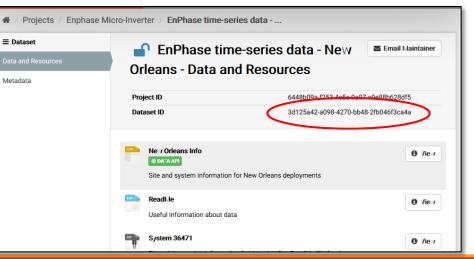
All items stored within the Data Hub is assigned a Universally Unique Identifier (UUID); projects, datasets, data files... **and even registered users** 

	A / Users / MCCat	
u <b>ser UUID</b> ( <b>a.k.a API Key)</b> – User Page	<b>Mr. Cool Cat</b> Test subject in DuraMat	Another Test Dataset  Resource  Test upload point  PEG  PRVATE _ds1  Resource  This dataset has no description
	Followers Datasets 0 13 Edits 321	Compendium of photovoltaic degradation rates  Resource an excel file showing degradation found in published literature across multiple international locations and across PV technologies. See associated publication at  CSV
	Username mccat Email Private mrccol.cat@friskieU.edu	Sample Instrument Synthesis and Log File     S Resources This is a data set containing a single synthesis process
	Member Since August 3, 2017 State active	RRIVATE Sample File Uploads     A series of different files showing different formats and examples     XML CSY JSON PDF
	API Key Private 55653878-2425-3878-8cf3- 96012f25505	General Standards Documentation 1 Resource Standards and best practices for researchers and users

#### 

🖀 / Projects / Enphas	e Micro-Inverter	
≡ Project	Enphase M	cro-Inverter
Datasets	Project ID	6448b09a-f253-4e5c-9a97-e9e88b628df5
Activity Stream		
		o-inverter data for PV systems located in Denver, Phoenix,
Edit Project	Minneapolis, Seattle, and New Orleans. E ID and include sequential data for all the	ata were provided as zipped CSV files with one file per PV system micro-inverters included in a PV system.
Bulk Edit Datasets		
Add Dataset	Project Metadata	
Members	Datasets	5
A Project Tree	Members	2
PV Field Data	Created On	2018-07-10T19:52:59.278123

#### 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.

ew Orleans Info	Email Maintainer	📕 Manage	Ownload	👗 Data API
e and system information for New Orleans deployments				
Data Explorer				
				<>> Embed
d Filter				
Grid Graph Map 46 records « 1 – 46	» Q	Search data	Go »	Filters
- lim Lic Py Lic Er	rst Data Collection: 2013-06- verter Family: M215 optitude: -89.91 / Module Model: RNG-250D tittude: 30.04 d: 21 uphase Site ID: 241036 verter Count: 24		<ul> <li>Latitude / Longit</li> <li>GeoJSON field</li> <li>Latitude field</li> <li>Latitude</li> <li>Longitude field</li> <li>Longitude</li> <li>Update</li> </ul>	vude fields
Laffer I Map ties by Stamen Desig			<ul> <li>Auto zoom to fea</li> <li>Cluster markers</li> </ul>	

CKAN Data API Access resource data via a web API with powerful query support. Further information in the main CKAN Data API and DataStore documentation. Endpoints » The Data API can be accessed via the following actions of the CKAN action API. Create https://datahub.duramat.org/api/3/action/datastore\_create Update / Insert https://datahub.duramat.org/api/3/action/datastore\_upsert Query https://datahub.duramat.org/api/3/action/datastore\_search Query (via https://datahub.duramat.org/api/3/action SQL) /datastore\_search\_sql Querying » Query example (first 5 results) https://datahub.duramat.org/api/3/action /datastore\_search?resource\_id=3796002f-017e-4420-95d3-89bfcf50bf91&limit=5 Query example (results containing 'jones') https://datahub.duramat.org/api/3/action/datastore\_search?q=jones& resource\_id=3796002f-017e-4420-95d3-89bfcf50bf91 Query example (via SQL statement) https://datahub.duramat.org/api/3/action/datastore\_search\_sql?sql=SELECT \* from "3796002f-017e-4420-95d3-89bfcf50bf91" WHERE title LIKE 'jones' Example: Javascript »

Example: Python »

import urllib url = 'https://datahub.duramat.org/api/3/action/datastore\_searc fileobj = urllib.urlopen(url) print fileobj.read()

### API Demo

Compare MAT NATIONAL RENEWABLE ENERGY LABORATORY • SANDIA NATIONAL LABORATORIES • LAWRENCE BERKELEY NATIONAL LABORATORY • SLAC NATIONAL ACCELERATOR LABORATORY

### Need more help?

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



# Thank you

-









