# ERDDAP: Easier access to scientific data

Rich Signell - USGS Filipe Fernandes - IOOS

OOI DDCI Kickoff Meeting, October 30, 2018





## What is ERDDAP?

A data brokerage service, reading from many different types of files, databases and services, and providing access via a single standardized interface

RESTful API for access in scientific analysis packages (Matlab, Python, R), web application developers (JavaScript), and by numerical modelers (Fortran, Bash)

Advanced search built-in, and also generates ISO and json-ld metadata records to allow search via sites like data.gov, and Google Dataset Search.

Widely used for delivery of "FAIR" data in the geoscience community (more than 50 server deployments worldwide)

# **USGS Integrated Ocean Observing System**



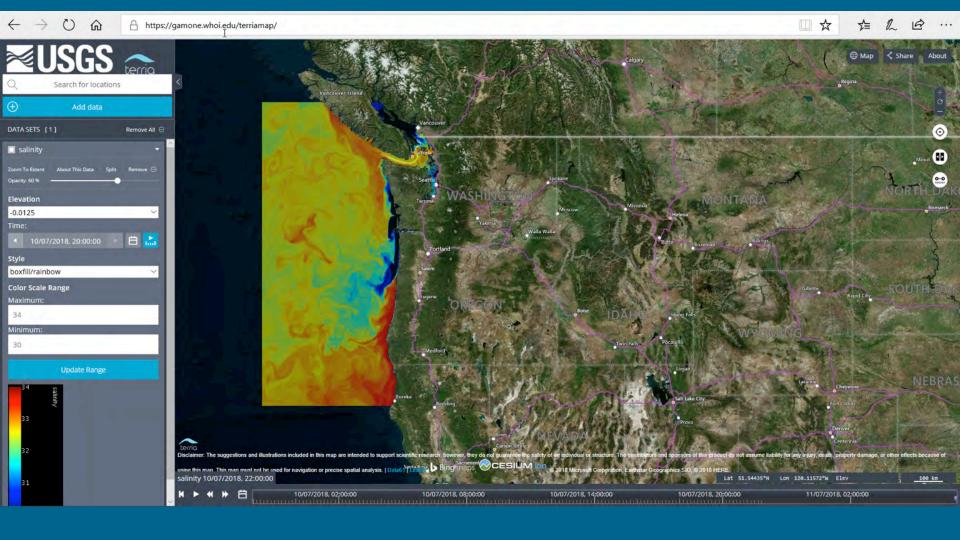
11 Regional Associations

17 Federal Partners

Web Standards:

Grid: OPeNDAP Sensor: ERDDAP

Image: WMS Metadata: CSW



#### ERDDAP > Advanced Search @

**Directions:** Specify as many or as few search criteria as you want, then click Search. Only the datasets that match **all** of the search criteria will appear in the results.

#### Full Text Search for Datasets @

#### 

cdm\_data\_type = (ANY) \( \frac{1}{2} \)

ioos\_category = (ANY) \( \frac{1}{2} \)
keywords = (ANY)

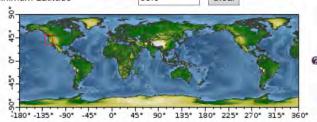
= (ANY)

= sea\_water\_practical\_salinity = (ANY)

## Search for Datasets that have Data within Longitude, Latitude, and Time Ranges Maximum Latitude = 50.0

Min and Max Longitude = -132.0 -117.0

Minimum Latitude = 33.0 Clear



Minimum Time = 2018-07-01 Maximum Time = 2018-07-15

long\_name standard name

variableName



#### Advanced Search Results

42 matching datasets, listed in alphabetical order.

Grid	Sub-	Table	Make	33	Source	247	Sum-	FGDC,	Back-		E		B. C. A. D.
DAP Data	set	DAP Data	A Graph	M S	Data Files	Title	mary	ISO, Metadata	ground Info	RSS	mail	Institution	Dataset ID
	set	data	graph		files	CE01ISSM BUOY CTDBP3	0	FIM	background ₽	RSS	$\bowtie$	Coastal and Globa 2	CE01ISSM-BUOY-001-CTDBP-FLORT
	set	data	graph		files	CE01ISSM MFN CTDBP2	0	FIM	background &	₹ RSS	$\bowtie$	Coastal and Globa 2	CE01ISSM-MFN-001-CTDBP-DOSTA
	set	data	graph		files	CE01ISSM NSIF CTDBP1	0	FIM	background ₽	₹ RSS	$\bowtie$	Coastal and Globa @	CE01ISSM-NSIF-001-CTDBP-DOSTA
	set	data	graph		files	CE01ISSM NSIF FLORT	0	FIM	background &	₹ RSS		Coastal and Globa @	CE01ISSM-NSIF-001-FLORT
	set	data	graph		files	CE01ISSM NSIF SUNA	0	FIM	background &	RSS	$\bowtie$	Coastal and Globa 2	CE01ISSM-NSIF-002-NUTNR
	set	data	graph		files	CE01ISSP CSPP CTDPF Recovered	0	FIM	background @	₹ RSS	$\bowtie$	Coastal and Globa 2	CE01ISSP-CSPP-002-CTDPF
	set	data	graph		files	CE01ISSP CSPP CTDPF Telemetry	0	FIM	background &	A RSS	$\bowtie$	Coastal and Globa @	CE01ISSP-CSPP-001-CTDPF
	set	data	graph		files	CE01ISSP CSPP FLORT Recovered	0	FIM	background &	₹ RSS		Coastal and Globa @	CE01ISSP-CSPP-002-FLORT
	set	data	graph		files	CE01ISSP CSPP FLORT Telemetry	0	FIM	background &	RSS	$\bowtie$	Coastal and Globa 2	CE01ISSP-CSPP-001-FLORT
	set	data	graph		files	CE01ISSP CSPP NUTNR Recovered	0	FIM	background &	M RSS	$\bowtie$	Coastal and Globa 2	CE01ISSP-CSPP-002-NUTNR
	set	data	graph		files	CE01ISSP CSPP OPTAA Recovered	0	FIM	background &	A RSS	$\bowtie$	Coastal and Globa @	CE01ISSP-CSPP-002-OPTAA
	set	data	graph		files	CE02SHSM BUOY METBK	0	F I M	background &	₹ RSS	$\boxtimes$	Coastal and Globa @	CE02SHSM-BUOY-001-METBK
	set	data	graph		files	CE02SHSM NSIF CTDBP	0	FIM	background &	RSS	$\bowtie$	Coastal and Globa 2	CE02SHSM-NSIF-001-CTDBP
	set	data	graph		files	CE02SHSM NSIF FLORT	0	FIM	background @	₹ RSS	$\bowtie$	Coastal and Globa @	CE02SHSM-NSIF-001-FLORT
	set	data	graph		files	CE02SHSM NSIF NUTNR	0	FIM	background &	₹ RSS	$\bowtie$	Coastal and Globa @	CE02SHSM-NSIF-001-NUTNR
	set	data	graph		files	CE02SHSP CSPP NUTNR Recovered	0	FIM	background &	₹ RSS	$\bowtie$	Coastal and Globa @	CE02SHSP-CSPP-002-NUTNR
	set	data	graph		files	CE04OSSM BUOY METBK	0	FIM	background 🗗	RSS	$\bowtie$	Coastal and Globa 2	CE04OSSM-BUOY-001-METBK
	set	data	graph		files	CE04OSSM NSIF CTDBP	0	FIM	background &	₹ RSS	$\bowtie$	Coastal and Globa 2	CE04OSSM-NSIF-001-CTDBP
	set	data	graph		files	CE04OSSM NSIF FLORT	0	FIM	background &	₹ RSS		Coastal and Globa @	CE04OSSM-NSIF-001-FLORT
	set	data	graph		files	CE04OSSM NSIF NUTNR	0	FIM	background &	₹ RSS		Coastal and Globa @	CE04OSSM-NSIF-001-NUTNR
	set	data	graph		files	CE06ISSM BUOY CTDBP3	0	FIM	background 🗗	R RSS	$\bowtie$	Coastal and Globa 2	CE06ISSM-BUOY-001-CTDBP-FLORT
	set	data	graph		files	CE06ISSM MFN CTDBP2	0	FIM	background &	₹ RSS	$\bowtie$	Coastal and Globa 2	CE06ISSM-MFN-001-CTDBP-DOSTA

#### ERDDAP > tabledap > Make A Graph @

Redraw the Graph (Please be patient. It may take a while to get the data.)

Dataset Title: CE06ISSP CSPP CTDPF Telemetry 

MRSS Institution: Coastal and Global Scales Nodes (CGSN) (Dataset ID: CE06ISSP-CSPP-001-CTDPF) Range: longitude = -124.27258 to -124.272°E, latitude = 47.133 to 47.13382°N, depth = 29.0 to 29.0m, time = 2017-07-26T07:45:49Z to 2018-07-20T07:58:35Z Information: Summary [9] License [9] FGDC | ISO 19115 | Metadata | Background [9] | Subset | Data Access Form | Files Time range: 14 v day(s) Graph Type: markers X Axis: time Y Axis pressure salinity Color: Optional Optional Constraints @ Water Pressure (dbar) Constraint #1 @ Constraint #2 @ 2018-07-01T00:00:00Z 2018-07-15T00:00:00Z time <= v >= >= <= . <= >= >= \ <= v Server-side Functions @ ~ @ (" **Graph Settings** 08 10 2018-07 10 ~ Marker Type: Square Size: 000000000000000 Color: Rainbow Color Bar: Continuity: Scale: 31.5 32 32.5 33 33.5 31 Maximum: 35 Practical Salinity (PSU)
CEOGISSP CSPP CTOPF Telemetry
Data courtesy of Coastal and Global Scales Nodes (CGSN) Minimum: N Sections: Y Axis Minimum: Maximum: descending Ascending:

#### ERDDAP > tabledap > Data Access Form @

10

Dataset Title: CE06ISSP CSPP CTDPF Telemetry 🖂 🔤

Institution: Coastal and Global Scales Nodes (CGSN) (Dataset ID: CE06ISSP-CSPP-001-CTDPF)

Information: Summary @ | License @ | FGDC | ISO 19115 | Metadata | Background & | Subset | Files | Make a graph

<= v	-2147483647	04 47 4000 47
		-2147483647
<= v	· · · · · · · · · · · · · · · · · · ·	
01T00:00:00Z <= \sqrt{2018-07-15T00:00:00}	DZ 2017-07-26T07:45:49Z	2018-07-20T07:58:35Z
<= v	~-+ 0	
<= v	~-+ <b>0</b>	
<= v	· -+ 0	
<= v	2.84188	4.042
<= v	0.703	26.208
<= v	22.7776	33.9715
<= v	7.7724	16.9595
<= v	1016.81170870495	1026.58764901618
<= v	v - + 0	
<= v	v - + 0	
	<= v  <= v	

File type: (more info)

.htmlTable - View a UTF-8 .html web page with the data in a table. Times are ISO 8601 strings.

Just generate the URL:

(Documentation / Bypass this form @)



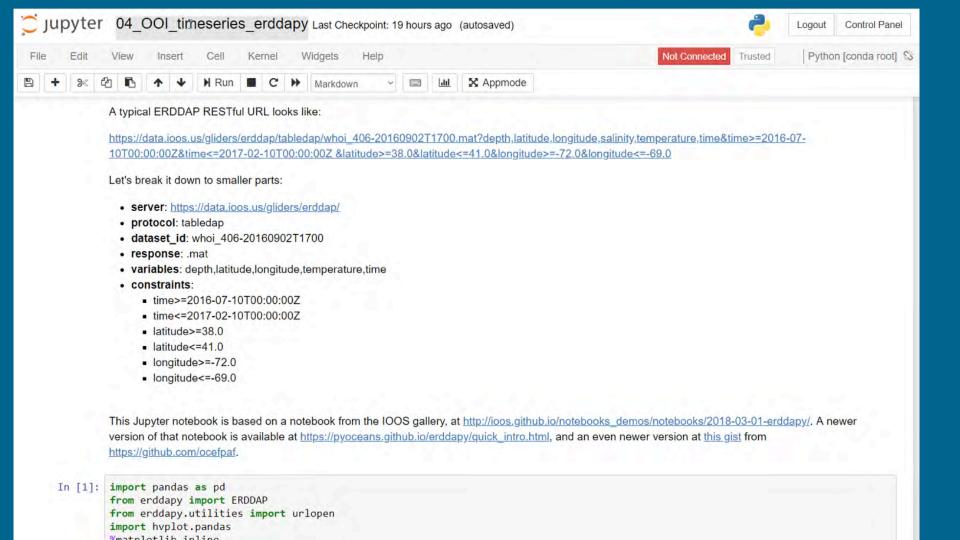
(Documentation / Bypass this form @)

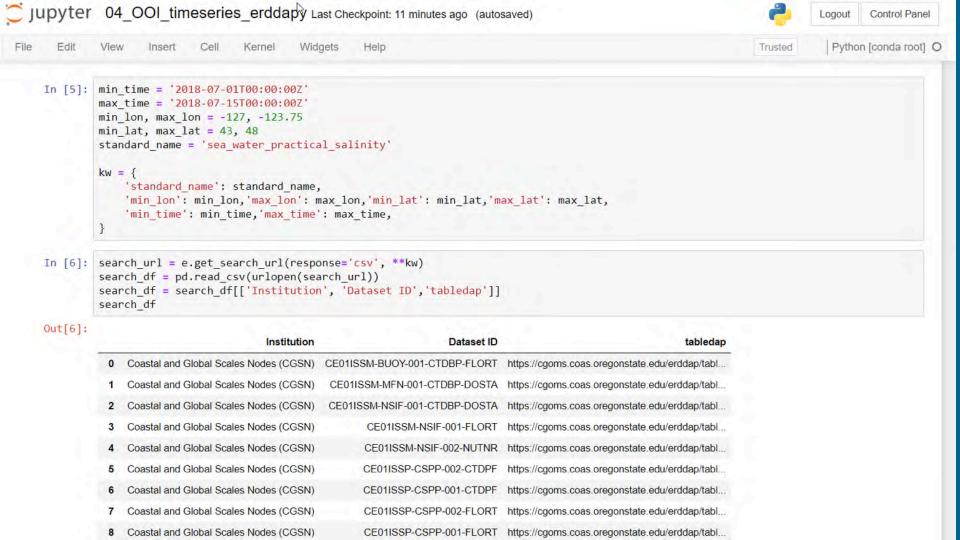
#### ERDDAP > tabledap > Data Access Form •

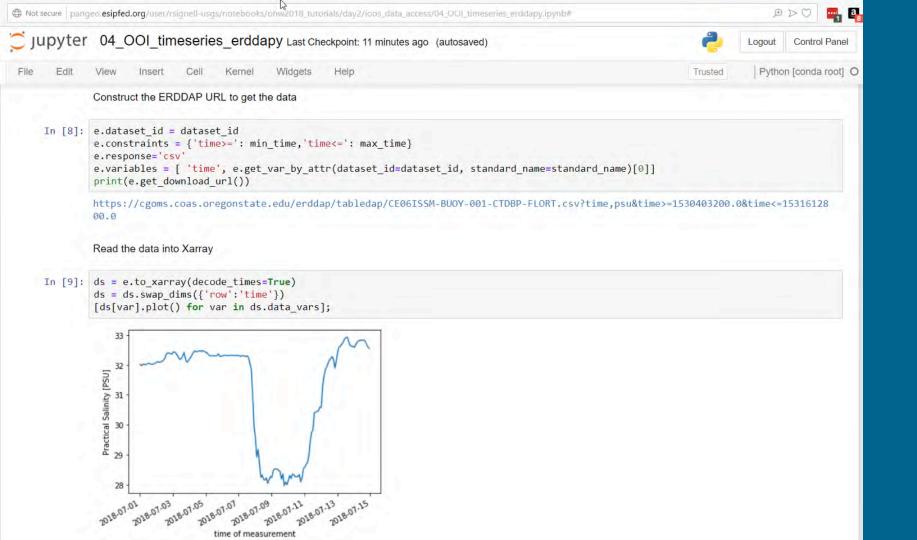
Dataset Title: CE06ISSP CSPP CTDPF Telemetry M 19853 Coastal and Global Scales Nodes (CGSN) (Dataset ID: CE06ISSP-CSPP-001-CTDPF) Information: Summary @ | License @ | FGDC | ISO 19115 | Metadata | Background @ | Subset | Files | Make a graph Optional Optional Minimum @ Maximum @ Variable Check All Uncheck All Constraint #1 @ Constraint #2 @ or a List of Values @ Crs @ -2147483647 -2147483647 >= v <= v station (station identifier) @ V - + 0 >= 4 <= 4 ☑ time (UTC) ❷ 2018-07-01T00:00:00Z 2018-07-15T00:00:00Z 2017-07-26T07:45:49Z 2018-07-20T07:58:35Z ☐ latitude (degrees\_north) ❷ >= 4 <= 4 · - + 0 v - + 0 asc - View OPeNDAP-style ISO-8859-1 comma-separated text. .csy - Download a ISO-8859-1 comma-separated text table (line 1: names; line 2: units; ISO 8601 times). .csvp - Download a ISO-8859-1 .csv file with line 1: name (units). Times are ISO 8601 strings v - + 0 .csv0 - Download a ISO-8859-1 .csv file without column names or units. Times are ISO 8601 strings. das - View the dataset's metadata via an ISO-8859-1 OPeNDAP Dataset Attribute Structure (DAS). 2.84188 4.042 .dds - View the dataset's structure via an ISO-8859-1 OPeNDAP Dataset Descriptor Structure (DDS). 0.703 26.208 .dods - OPeNDAP clients use this to download the data in the DODS binary format. .esriCsv - Download a ISO 8859 1 .csv file for ESRI's ArcGIS 9.x and below (separate date and time columns). 22.7776 33.9715 fgdc - View the dataset's UTF-8 FGDC .xml metadata. 7.7724 16 9595 .geoJson - Download longitude.latitude.otherColumns data as a UTF-8 GeoJSON .ison file. 1016.81170870495 1026 58764901618 graph - View a Make A Graph web page. ~ - + 0 help - View a web page with a description of tabledap V - + 0 .html - View an OPeNDAP-style HTML Data Access Form. .htmlTable - View a UTF-8 .html web page with the data in a table. Times are ISO 8601 strings. iso19115 - View the dataset's ISO 19115-2/19139 UTF-8 .xml metadata. .itx - Download an ISO-8859-1 Igor Text File. Each response column becomes a wave json - View a table-like UTF-8 JSON file (missing value = 'null'; times are ISO 8601 strings). isonICSV - View a UTF-8 JSON Lines CSV file (missing value = 'null': times are ISO 8601 strings) isonIKVP - View a UTF-8 JSON Lines file with Key: Value pairs (missing value = 'null'; times are ISO 8601 strings). mat - Download a MATLAB binary file. .htmlTable - View a UTF-8 .html web page with the data in a table. Times are ISO 8601 strings. Just generate the URL:

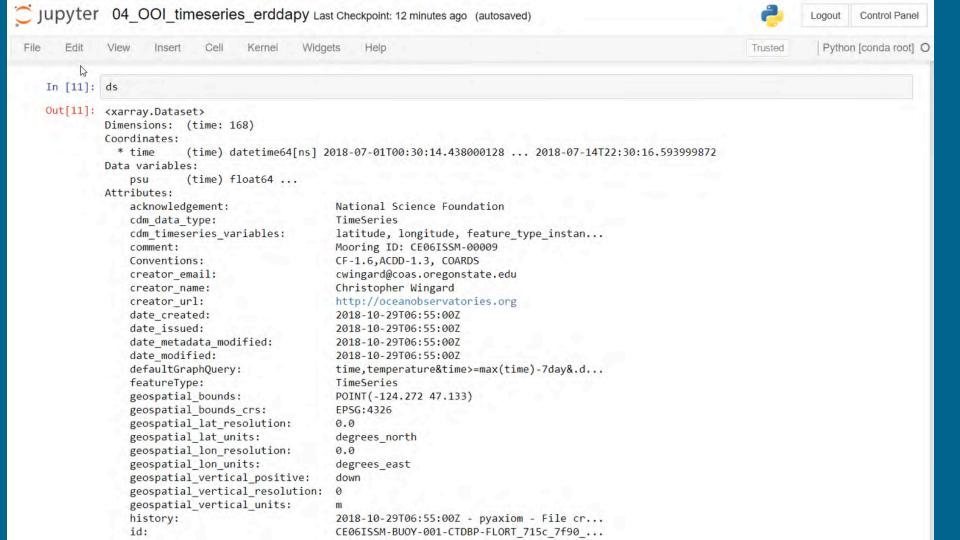


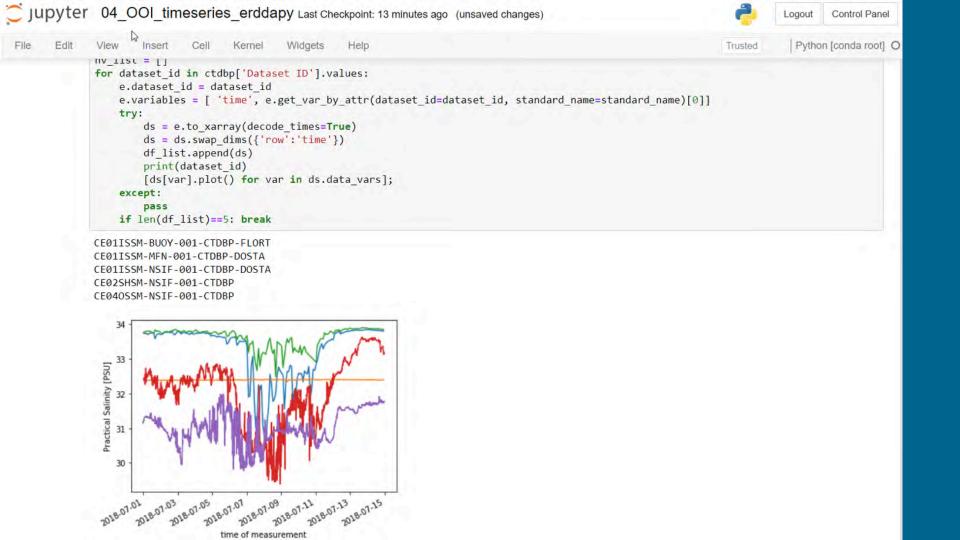
data services for machine-to-machine access. **ERDDAP** (see <a href="here">here</a> and <a href="here">here</a> and <a href="here">here</a>) is one of those applications. "ERDDAP is a data server that gives you a simple, consistent way to download data in the format and the spatial and temporal coverage that you want. ERDDAP is a web application with an interface for people to use. It is also a RESTful web service that allows data access directly from any computer program (e.g. Matlab, R, or webpages)."

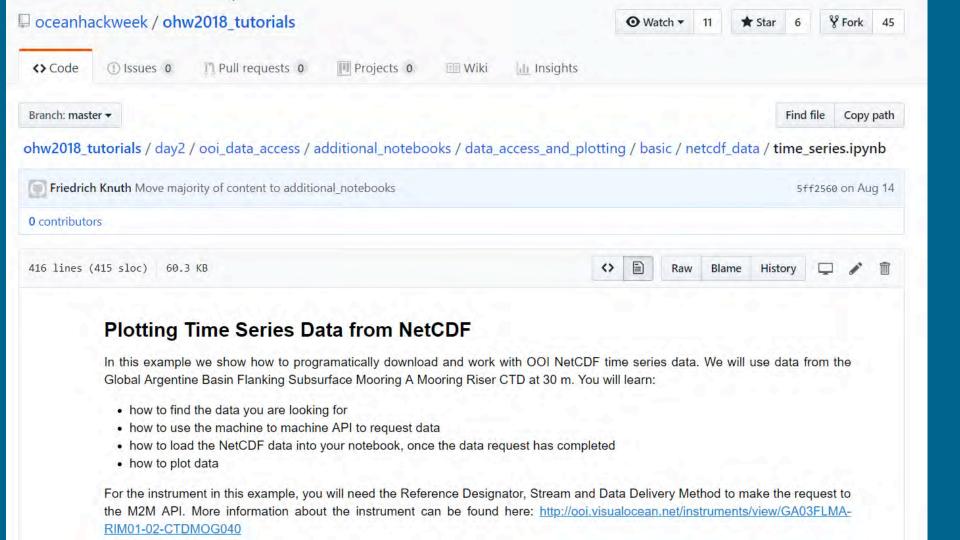


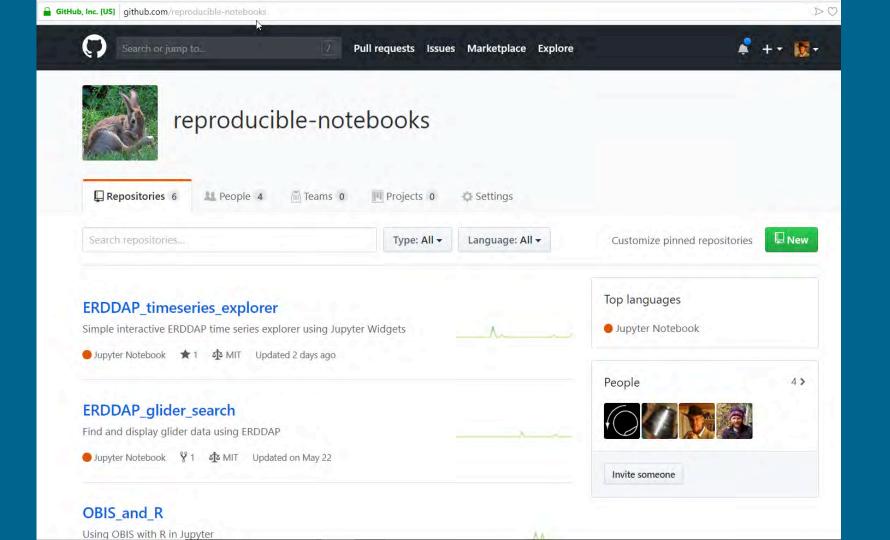












## **Explore ERDDAP timeseries data using Jupyter Widgets**

Inspired by <u>Jason Grout's excellent ESIP Tech Dive talk on "Jupyter Widgets"</u>, this notebook uses the ipyleaflet and bqplot widgets to interactively explore the last two weeks of time series data from an ERDDAP Server. Select a standard\_name from the list, then click a station to see the time series.

NOTE: To access a protected ERDDAP endpoint is protected, you can add a ~/.netrc file like this:

```
machine cgoms.coas.oregonstate.edu
login <username>
password <password>
```

```
In [1]: import numpy as np import pandas as pd
```

In [2]:

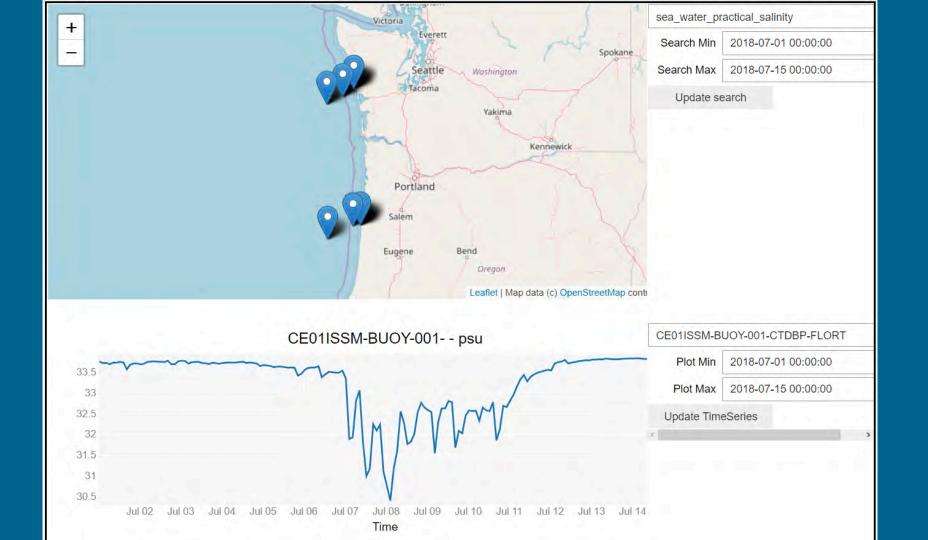
```
import pendulum
```

ipyleaflet and bqplot are both Jupyter widgets, so can interact with Python like any other widget. Since we want to click on a map in a notebook and get an interactive time series plot, they are perfect tools to use here.

```
In [3]: import ipyleaflet as ipyl
import bqplot as bq
import ipywidgets as ipyw
```

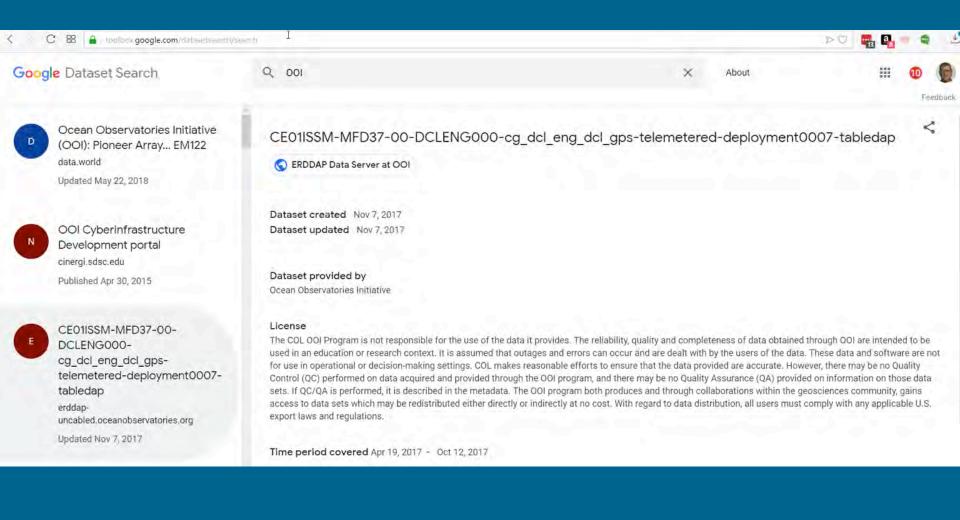
To make working with ERDDAP simpler, we use enddapy, a high-level python interface to ERDDAP's RESTful API

```
In [4]: from erddapy import ERDDAP
    from erddapy.utilities import urlopen
```



## **CGSN Dashboard**













## ${\tt ERDDAP > info > CE01ISSM-MFD37-00-DCLENG000-cg\_dcl\_eng\_dcl\_gps-telemetered-deployment0007-tabledapaled and the compact of the compact of$

Grid DAP Data	CAT			M	Source Data Files	Title	Sum- mary	FGDC, ISO, Metadata	Back- ground Info	RSS	E mail	Institution	
		data	graph			Coastal Endurance Oregon Inshore Surface Mooring Seafloor Multi-Function Node (MFN) Data Concentrator Logger (DCL) cg_dcl_eng_dcl_gps - Deployment 0007 (telemetered)	0	FIM	background ₽	₹ RSS	$\boxtimes$	001	CE01ISSM-MFD37-00-DCLENG000-cg_dc

#### The Dataset's Variables and Attributes

Row Type	Variable Name	Attribute Name	Data Type	Value
attribute	NC_GLOBAL	cdm_data_type	String	Point
attribute	NC_GLOBAL	collection_method	String	telemetered
attribute	NC_GLOBAL	Conventions	String	CF-1.6, COARDS, ACDD-1.3, NCCSV-1.0
attribute	NC_GLOBAL	creator_name	String	Ocean Observatories Initiative
attribute	NC_GLOBAL	creator_url	String	http://oceanobservatories.org/@
attribute	NC_GLOBAL	date_created	String	2017-11-07T19:47:21.731626
attribute	NC_GLOBAL	date_modified	String	2017-11-07T19:47:21.731631
attribute	NC_GLOBAL	Easternmost_Easting	double	-124.0954
attribute	NC_GLOBAL	featureType	String	Point
attribute	NC_GLOBAL	geospatial_lat_max	double	44.65692
attribute	NC_GLOBAL	geospatial_lat_min	double	44.65665
attribute	NC_GLOBAL	geospatial_lat_resolution	double	0.1
attribute	NC_GLOBAL	geospatial_lat_units	String	degrees_north
attribute	NC_GLOBAL	geospatial_lon_max	double	-124.0954

# **ERDDAP vs M2M**

	ERDDAP	M2M			
RESTful Interface	Yes	Yes			
Widely used by community	Yes	No			
Advanced Search	Yes	No			
ISO, JSON-LD Metadata	Yes	No			
Output types	40 (NetCDF, json, mat, csv)	2 (NetCDF, json)			

# Summary

- ERDDAP allows easier access to OOI Data
- ERDDAP makes OOI "FAIR" (Findable, Accessible, Interoperable, Reusable)
- ERDDAP works with the existing system
- Easier access to OOI Data results in more use by researchers
- Easier access to OOI Data allows more efficient data analysis, leaving more time for actual science
- Let's make ERDDAP a first class citizen in OOI
- Let's build OOI science and end-user applications using ERDDAP as a backend for data search and access