OCEAN OBSERVATORIES INITIATIVE

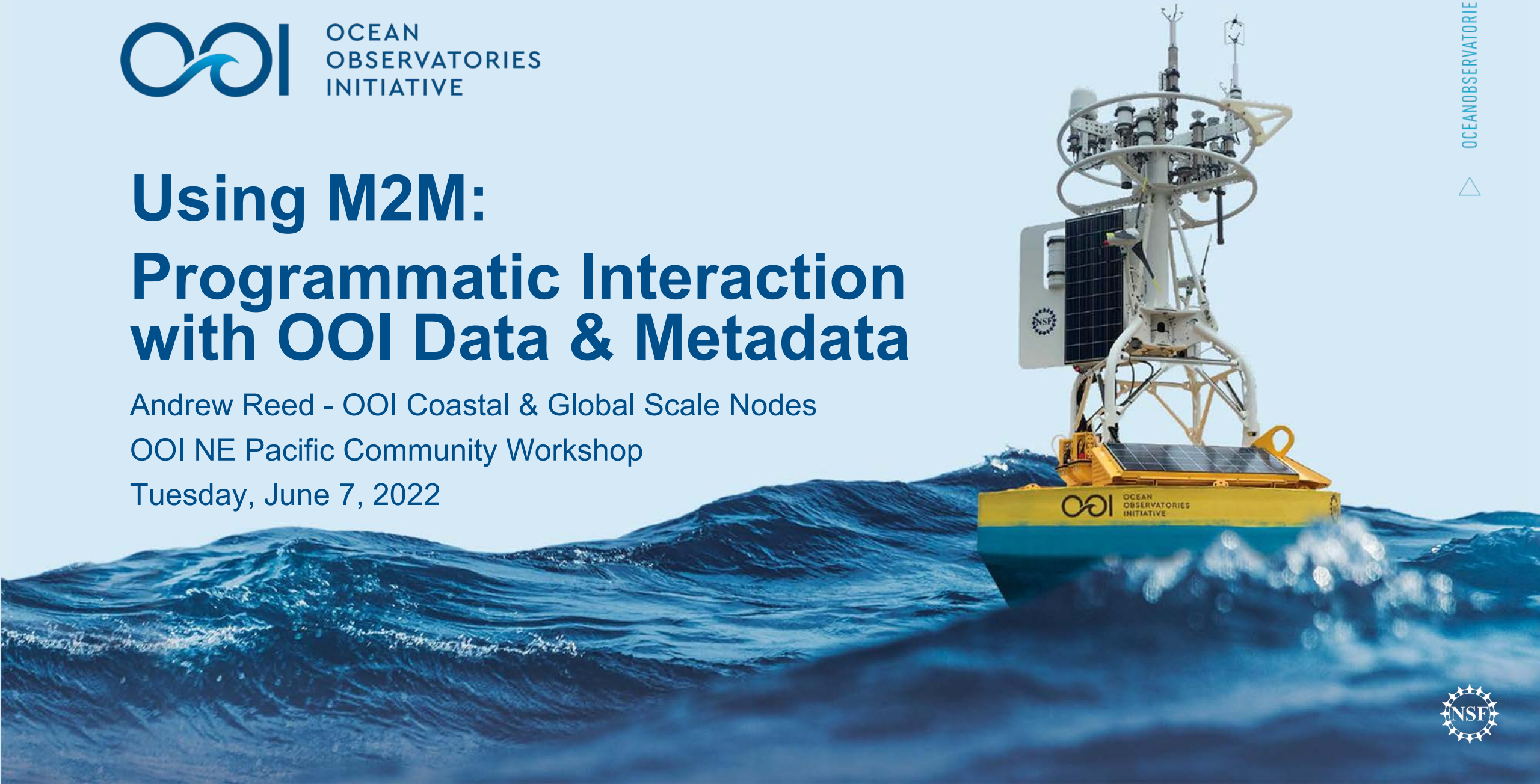
OCEANOBSERVATORIES.ORG

Using M2M:
Programmatic Interaction with OOI Data & Metadata

Andrew Reed - OOI Coastal & Global Scale Nodes

OOI NE Pacific Community Workshop

Tuesday, June 7, 2022



NSF

OOI's API: Machine-to-Machine (M2M)

- What is M2M?
  - M2M is the name for OOI's Application Programming Interface (API)
  - M2M allows for the programmatic access (i.e. scripting) to OOI's data and metadata
  - Allows for download of data in **JSON, netCDF,** and **csv** formats from OOI's THREDDS server
  - Language agnostic:
    - Examples principally utilize either Python or MatLab
    - Developing user base for sharing of code and analysis (https://github.com/oceanobservatories/ooi-data-explorations)

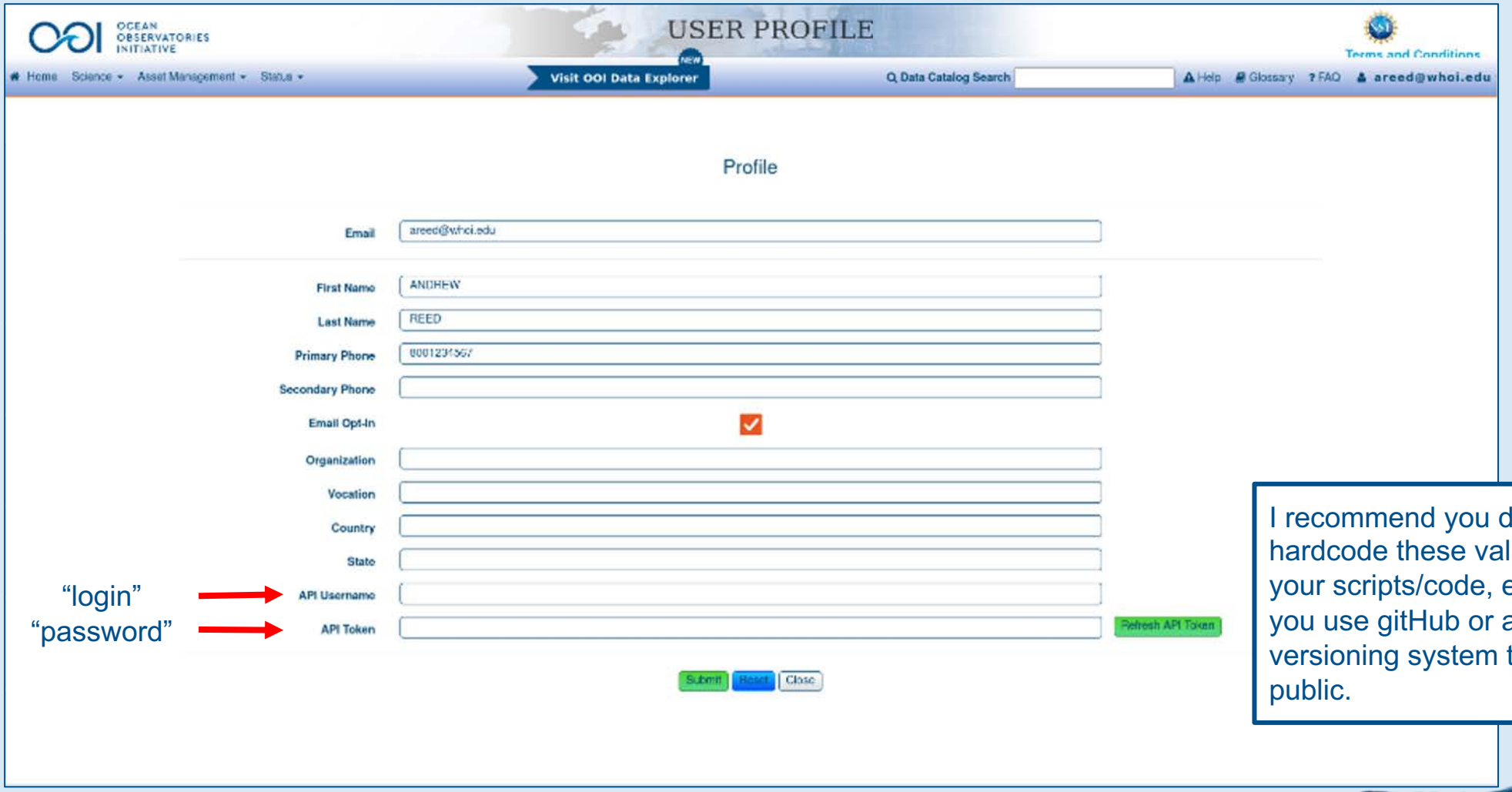
- What is accessible?
  - Metadata
  - Calibrations
  - Deployment Information
  - Data
    - Fast **synchronous** requests returned as a **JSON** object but limited data (<20,000 data points)
    - Slower **asynchronous** requests returned as either **JSON**, **CSV**, or **netCDF** with no data point limit
  - Annotations

- Key Handouts
  - API Cheat Sheet
  - OOI Decoder Sheet
  - OOI Instrument List

Terminology (Handout: Decoder)

- Array
  - One of 5 major research components that make up OOI (e.g. Ocean Station Papa or Coastal Endurance)
- Site
  - A specific geographic location within an array (e.g. Global Ocean Station Papa Flanking Mooring A = GP03FLMA)
- Platform
  - A set of infrastructure within an array that host a complement of integrated scientific instrument. May be fixed (mooring) or mobile (profilers or gliders)
- Node
  - A section of a platform with one or more computers and power controllers. Instruments on platforms are plugged into nodes. (e.g. mooring riser on Global Ocean Station Papa Flanking Mooring A = GP03FLMA-RIS01)
- Instrument/Sensor
  - Terms often used interchangeably. An **instrument** is a piece of equipment used to collect data. A **sensor** is a part of an instrument which measures a specific quantity.
  - Each instrument has a **Unique ID** (**UID)**
- Reference Designator
  - Identifies a particular instrument on a particular node/platform at a particular site.
  - Example: **GP03FLMA-RIS01-02-DOSTAD000** = Global Ocean Station Papa Flanking Mooring A - Mooring Riser - Dissolved Oxygen Sensor
- Method
  - **Telemetered:** Data returned wirelessly. Often truncated or decimated due to size.
  - **Recovered_host/Recovered_inst**: Data downloaded directly from either the mooring computer or from the instrument.
  - **Streamed:** Data accessible nearly instantaneously. Limited to cabled array and Endurance shelf stations
- Data Stream
  - Raw data feed from a sensor that has been read, parsed, and separated based on content (e.g. engineering, science, metadata, etc.)
- Parameters
  - Also sometimes called data variables.
  - A particular value returned from a sensor (e.g. practical salinity from a CTD). Many parameters are in a given datastream.
- Annotation
  - A human-in-the-loop assessment of the data added to the datasets for specified time periods by the operators.
  - Provides info on data quality, performance, issues, and context.

Setup & Requirements



"login" →
"password" →

I recommend you do **NOT** hardcode these values into your scripts/code, especially if you use gitHub or another versioning system that is public.

Data/Metadata Access (Handout: API Cheat Sheet)

Category	Access point	Description
Deployment	12587/events/deployment/inv/	Access deployment numbers as well as the asset & calibration info for specified instrument & deployment, and deployment times/cruises
	12587/asset/deployments	Asset & calibration info for all deployment for the specified UID
Calibration	12587/asset/cal?uid= OR ?assetid=	Return all calibration info for a given uid or assetId
	12587/asset/cal?refdes=	Return list of deployments with calibrations for a given reference designator
Asset	12587/asset?uid= OR ?serialnumber=	Asset information by unique id or instrument serial number
Preload	12575/parameter/	Retrieve information for a parameter (i.e. variable) given its ID number
	12575/stream/byname/	Retrieve information for a stream given its name
Annotations	/12580/anno/find?=	Retrieve annotations for a specific time period and for a given reference designator (optional: stream and method)
Vocab	12586/vocab/inv/	Get the vocabulary (descriptions) for a sensor
Data	12576/sensor/inv/	Can access the data from OOI using either a **synchronous** (returns JSON; limited to 20000 data points) or **asynchronous** (returns netCDF, CSV, or JSON; not data limit)

Finding Data: Step 1

Endpoint: **m2m/12576/sensor/inv**

```
API Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12576/sensor/inv

Returns the following list of sites:

['CE01ISSM', 'CE01ISSP', 'CE02SHBP', 'CE02SHSM', 'CE02SHSP', 'CE04OSBP', 'CE04OSPD', 'CE04OSPI', 'CE04OSPS', 'CE04OS
SM', 'CE05MOAS', 'CE06ISSM', 'CE06ISSP', 'CE07SHSM', 'CE07SHSP', 'CE09OSPM', 'CE09OSSM', 'CP01CNPM', 'CP01CNSM', 'CP
01CNSP', 'CP02PMCI', 'CP02PMCO', 'CP02PMUI', 'CP02PMUO', 'CP03ISPM', 'CP03ISSM', 'CP03ISSP', 'CP04OSPM', 'CP04OSSM',
'CP05MOAS', 'GA01SUMO', 'GA02HYPM', 'GA03FLMA', 'GA03FLMB', 'GA05MOAS', 'GI01SUMO', 'GI02HYPM', 'GI03FLMA', 'GI03FLM
B', 'GI05MOAS', 'GP02HYPM', 'GP03FLMA', 'GP03FLMB', 'GP05MOAS', 'GS01SUMO', 'GS02HYPM', 'GS03FLMA', 'GS03FLMB', 'GS0
5MOAS', 'RS01OSBP', 'RS01SBPD', 'RS01SBPS', 'RS01SHBP', 'RS01SHDR', 'RS01SLBS', 'RS01SUM1', 'RS01SUM2', 'RS03ASHS',
'RS03AXBS', 'RS03AXPD', 'RS03AXPS', 'RS03AXSM', 'RS03CCAL', 'RS03ECAL', 'RS03INT1', 'RS03INT2', 'SSRSPACC']
```

Select a site, e.g. Global Ocean Station Papa Flanking Mooring A (GP03FLMA)

Finding Data: Step 2

Endpoint: **m2m/12576/sensor/inv/GP03FLMA**

```
API Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12576/sensor/inv/GP03FLMA

Returns the following list of nodes on GP03FLMA:

['RIM01', 'RIS01']
```

Returns two nodes: the mooring sphere (RIS01) and the riser (RIM01)

Select a node: **Global Ocean Station Papa Flanking Mooring A - Mooring Sphere (GP03FLMA-RIS01)**

Finding Data: Step 3 (Handout: Instrument List)

Endpoint: **m2m/12576/sensor/inv/GP03FLMA/RIS01**

```
API Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12576/sensor/inv/GP03FLMA/RIS01

Returns the following list of sensors on GP03FLMA-RIS01:

['00-SIOENG000', '03-DOSTAD000', '04-PHSENF000', '05-FLORTD000']
```

Returns four sensors:

- 00-SIOENG000 - engineering data w/no scientific interest
- 03-DOSTAD000 - dissolved oxygen sensor
- 04-PHSENF000 - pH sensor
- 05-FLORTD000 - fluorescense/chlorophyll/turbidity sensor

We're going to select the oxygen sensor '03-DOSTAD000'. This also allows us to construct the **Reference Designator:** GP03FLMA-RIS01-03-DOSTAD000

Finding Data: Step 4

Endpoint: **m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000**

```
API Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000

Returns the following list of data delivery methods on GP03FLMA-RIS01-03-DOSTAD000:

['recovered_host', 'telemetered']
```

Returns the data delivery methods:

- **recovered_host**: this is data recorded by the mooring computer and downloaded after recovering the mooring
- **telemetered**: this is data that was transmitted from the mooring back-to-shore. For some instruments it is truncated or decimated.

Finding Data: Step 5

Endpoint: **m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000/recovered_host**

```
API Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000/recovered_h
ost

Returns the following list of methods for GP03FLMA-RIS01-03-DOSTAD000 recovered_host:

['dosta_abcdjm_sio_instrument_recovered', 'dosta_abcdjm_sio_metadata_recovered']
```

Returns the data streams:

- **dosta_abcdjm_sio_instrument_recovered -** this stream has science data
- **dosta_abcdjm_sio_metadata_recovered -** this stream has metadata

Now we can go ahead and request data!

Full data request: **m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000/recovered_host/dosta_abcdjm_sio_instrument_recovered**

Requesting Data: Synchronous

Synchronous data request

- Faster
- Returned as JSON
- Request Specifications
  - limit (required): specifies number of data points with a maximum of 20000
  - beginDT (optional): start date as YYYY-mm-ddTHH:MM:SS.fffZ format
  - endDT (optional): end date in same format as beginDT
  - parameters (optional): numeric IDs of which parameters to get

Requesting Data: Synchronous

Building a request

- Want dissolved oxygen data for the year of 2015 from Ocean Station Papa Flanking Mooring A oxygen sensor
- Specifications
  - limit: 20000
  - beginDT: 2015-01-01T00:00:00.000Z
  - endDT: 2016-01-01T00:00:00.000Z
  - parameters: 7 (time), 14 (dissolved oxygen)
- Example Request
  - **m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000/recovered_host/dosta_abcdjm_sio_instrument_recovered?limit=20000&beginDT=2015-01-01T00:00:00.000Z&endDT=2016-01-01T00:00:00.000Z&parameters=7,14**
  - **IMPORTANT:** If you specify parameters, you must also specify time (“7”) or it will not be included!

Requesting Data: Synchronous

Step 1: Request the data

```
In [90]: # Ahhhh, No time was included! We forget to explicitly ask for it (parameter ID is 7).
method = "recovered_host"
stream = "dosta_abcdjm_sio_instrument_recovered"

# Request the oxygen data from the
data_url = "/".join((URLS["data"], site, node, sensor, method, stream))

params = {
    "beginDT": "2015-01-01T00:00:01.000Z",
    "endDT": "2016-01-01T00:00:01.000Z",
    "limit": "20000",
    "parameters": "7,14"
}

data = get_api(data_url, params)
data
```

Step 2: Data as JSON object

```
Out[90]: [{'flort_sample-pressure': None,
  'ctdmo_ghqr_sio_mule_instrument-ctdmo_seawater_temperature': 8.378609999999998,
  'dissolved_oxygen_qc_results': 29,
  'dissolved_oxygen': 281.74808247723314,
  'ctdmo_ghqr_sio_mule_instrument-practical_salinity': 32.46377730191175,
  'time': 3629059256.0,
  'dissolved_oxygen_qc_executed': 29,
  'pk': {'node': 'RIS01',
   'stream': 'dosta_abcdjm_sio_instrument_recovered',
   'subsite': 'GP03FLMA',
   'deployment': 2,
   'time': 3629059256.0,
   'sensor': '03-DOSTAD000',
   'method': 'recovered_host'}},
 {'flort_sample-pressure': None,
  'ctdmo_ghqr_sio_mule_instrument-ctdmo_seawater_temperature': 8.380299999999998,
  'dissolved_oxygen_qc_results': 29,
  'dissolved_oxygen': 281.6007067236025,
  'ctdmo_ghqr_sio_mule_instrument-practical_salinity': 32.46373369853948,
```

Desired parameter (dissolved oxygen)

Step 3: Parse JSON

	flort_sample-pressure	ctdmo_ghqr_sio_mule_instrument-ctdmo_seawater_temperature	dissolved_oxygen_qc_results	dissolved_oxygen	ctdmo_ghqr_sio_mule_instrument-practical_salinity	time	dissolved_oxygen_qc_executed	pk	deployment
0	None	[illegible]	[illegible]	[illegible]	[illegible]	2015-01-01 00:00:56	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
1	None	8.380300	29	281.600707	32.463734	2015-01-01 00:15:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
2	None	8.382800	29	281.374993	32.463828	2015-01-01 00:30:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
3	None	8.381400	29	281.228952	32.463125	2015-01-01 00:45:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
4	None	8.379800	29	280.712604	32.462216	2015-01-01 01:00:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
5	None	8.380700	29	280.231878	32.461662	2015-01-01 01:15:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
6	None	8.382200	29	280.279984	32.461787	2015-01-01 01:30:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
7	None	8.383600	29	279.829217	32.459774	2015-01-01 01:45:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
8	None	8.385447	29	279.692361	32.457720	2015-01-01 02:00:56	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2
9	None	8.387700	29	279.624252	32.457226	2015-01-01 02:15:01	29	{'node': 'RIS01', 'stream': 'dosta_abcdjm_sio_...	2

Data from co-located sensors needed in calculation of desired parameter

Requesting Data: Asynchronous

Asynchronous data request

- Slower
- Can return data as either JSON, CSV, or as netCDF (default)
- No data point limit
- Request Specifications
  - limit (required): if not specified, defaults netCDF
  - beginDT (optional): start date as YYYY-mm-ddTHH:MM:SS.fffZ format
  - endDT (optional): end date in same format as beginDT
  - parameters (optional): numeric IDs of which parameters to get
  - include_provenance (optional, default False): include a provenance file which specifies data processing paths
  - include_annotations (optional, default False): include a file with data annotations

Requesting Data: Asynchronous

Building a request

- Want dissolved oxygen data for the year of 2015 from Ocean Station Papa Flanking Mooring A oxygen sensor
- Specifications
  - beginDT: 2015-01-01T00:00:00.000Z
  - endDT: 2016-01-01T00:00:00.000Z
  - parameters: 7 (time), 14 (dissolved oxygen)
- Example Request
  - **m2m/12576/sensor/inv/GP03FLMA/RIS01/03-DOSTAD000/recovered_host/dosta_abcdjm_sio_instrument_recovered?beginDT=2015-01-01T00:00:00.000Z&endDT=2016-01-01T00:00:00.000Z&parameters=7,14**

Requesting Data: Asynchronous

Step 1: Get THREDDS server url from M2M


```
{'requestUUID': '5f7bfedc-afb9-485a-8751-0754f762f24a',
 'outputURL': 'https://opendap-west.oceanobservatories.org/thredds/catalog/ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/catalog.html',
 'allURLs': ['https://opendap-west.oceanobservatories.org/thredds/catalog/ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/catalog.html',
  'https://downloads-west.oceanobservatories.org/async_results/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered'],
 'sizeCalculation': 1950431,
 'timeCalculation': 60,
 'numberOfSubJobs': 14}
```


Step 2: Get dataset catalog using THREDDS server url


```
['catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIM01-02-CTDMOG040-recovered_host-ctdmo_ghqr_sio_mule_instrument_20150101T000001-20150604T230001.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150101T000056-20150125T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150126T000046-20150222T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150223T000035-20150322T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150323T000025-20150419T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150420T000015-20150517T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150518T000005-20150605T144501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0002_GP03FLMA-RIS01-05-FLORTD000-recovered_host-flort_sample_20141231T235501-20150605T144501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIM01-02-CTDMOG040-recovered_host-ctdmo_ghqr_sio_mule_instrument_20150606T224501-20160111T000001.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150606T224501-20150614T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150615T000001-20150712T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150713T000001-20150809T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150810T000002-20150906T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20150907T000001-20151004T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20151005T000001-20151101T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20151102T000001-20151129T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20151130T000001-20151227T234501.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered_20151228T000001-20160101T000001.nc',
 'catalog.html?dataset=ooi/areed@whoi.edu/20220523T171503214Z-GP03FLMA-RIS01-03-DOSTAD000-recovered_host-dosta_abcdjm_sio_instrument_recovered/deployment0003_GP03FLMA-RIS01-05-FLORTD000-recovered_host-flort_sample_20150606T224501-20160102T000001.nc']
```

Requesting Data: Asynchronous

See Chris Wingard's Breakout session
"Navigating THREDDS & ERDDAP"
for more info

Requesting Data: Asynchrono

Opened netCDF files as a dataset using xarray:

xarray.Dataset

- Dimensions: (**time**: 34914)
- Coordinates:

obs	(time)	int32	0 1 2 3 4 5 ... 380 381 382 383 384
time	(time)	datetime64[ns]	2015-01-01T00:00:56 ... 2016-01-...

- Data variables:

provenance	(time)	|S36	b'cbad28e4-547a-405e-b41e-7148a2...
deployment	(time)	int32	2 2 2 2 2 2 2 2 ... 3 3 3 3 3 3 3 3
id	(time)	|S36	b'f0a59d16-02fc-4990-a0e5-2f5887...
practical_salinity	(time)	float64	32.46 32.46 32.46 ... 32.54 32.54
ctdmo_seawate...	(time)	float64	8.379 8.38 8.383 ... 7.594 7.593
dissolved_oxygen	(time)	float64	281.7 281.6 281.4 ... 287.2 284.8
pressure	(time)	float64	nan nan nan nan ... nan nan nan nan
dissolved_oxyg...	(time)	uint8	29 29 29 29 29 ... 29 29 29 29 29
dissolved_oxyg...	(time)	uint8	29 29 29 29 29 ... 29 29 29 29 29

- Attributes: (70)

xarray.Dataset

- Dimensions: (**time**: 34914)
- Coordinates:

obs	(time)	int32	0 1 2 3 4 5 ... 380 381 382 383 384
time	(time)	datetime64[ns]	2015-01-01T00:00:56 ... 2016-01-...

- Data variables:

provenance	(time)	|S36	b'cbad28e4-547a-405e-b41e-7148a2...
deployment	(time)	int32	2 2 2 2 2 2 2 2 ... 3 3 3 3 3 3 3 3
id	(time)	|S36	b'f0a59d16-02fc-4990-a0e5-2f5887...
practical_salinity	(time)	float64	32.46 32.46 32.46 ... 32.54 32.54

comment : Salinity is generally defined as the concentration of dissolved salt in a parcel of seawater. Practical Salinity is a more specific unitless quantity calculated from the conductivity of seawater and adjusted for temperature and pressure. It is approximately equivalent to Absolute Salinity (the mass fraction of dissolved salt in seawater) but they are not interchangeable.

data_product_id... PRACSAL_L2
precision : 4
coordinates : time lat lon pressure
long_name : Practical Salinity
standard_name : sea_water_practical_salinity
units : 1
instrument : GP03FLMA-RIM01-02-CTDMOG040
stream : ctdmo_ghqr_sio_mule_instrument

ctdmo_seawate...	(time)	float64	8.379 8.38 8.383 ... 7.594 7.593

comment : Seawater temperature near the sensor.
data_product_id... TEMPWAT_L1
precision : 4
coordinates : time lat lon pressure
long_name : Seawater Temperature
standard_name : sea_water_temperature
units : °C
instrument : GP03FLMA-RIM01-02-CTDMOG040
stream : ctdmo_ghqr_sio_mule_instrument

dissolved_oxygen	(time)	float64	281.7 281.6 281.4 ... 287.2 284.8

comment : Dissolved Oxygen Concentration from the Stable Response Dissolved Oxygen Instrument is a measure of the concentration of gaseous oxygen mixed in seawater. This data product is corrected for salinity, temperature, and depth.

long_name : DO from Onboard Calculation - Corrected
precision : 4
coordinates : time lat lon pressure
data_product_id... DOXYGEN_L2
standard_name : moles_of_oxygen_per_unit_mass_in_sea_water
units : µmol kg-1
ancillary_variabl... practical_salinity pressure ctdmo_seawater_temperature

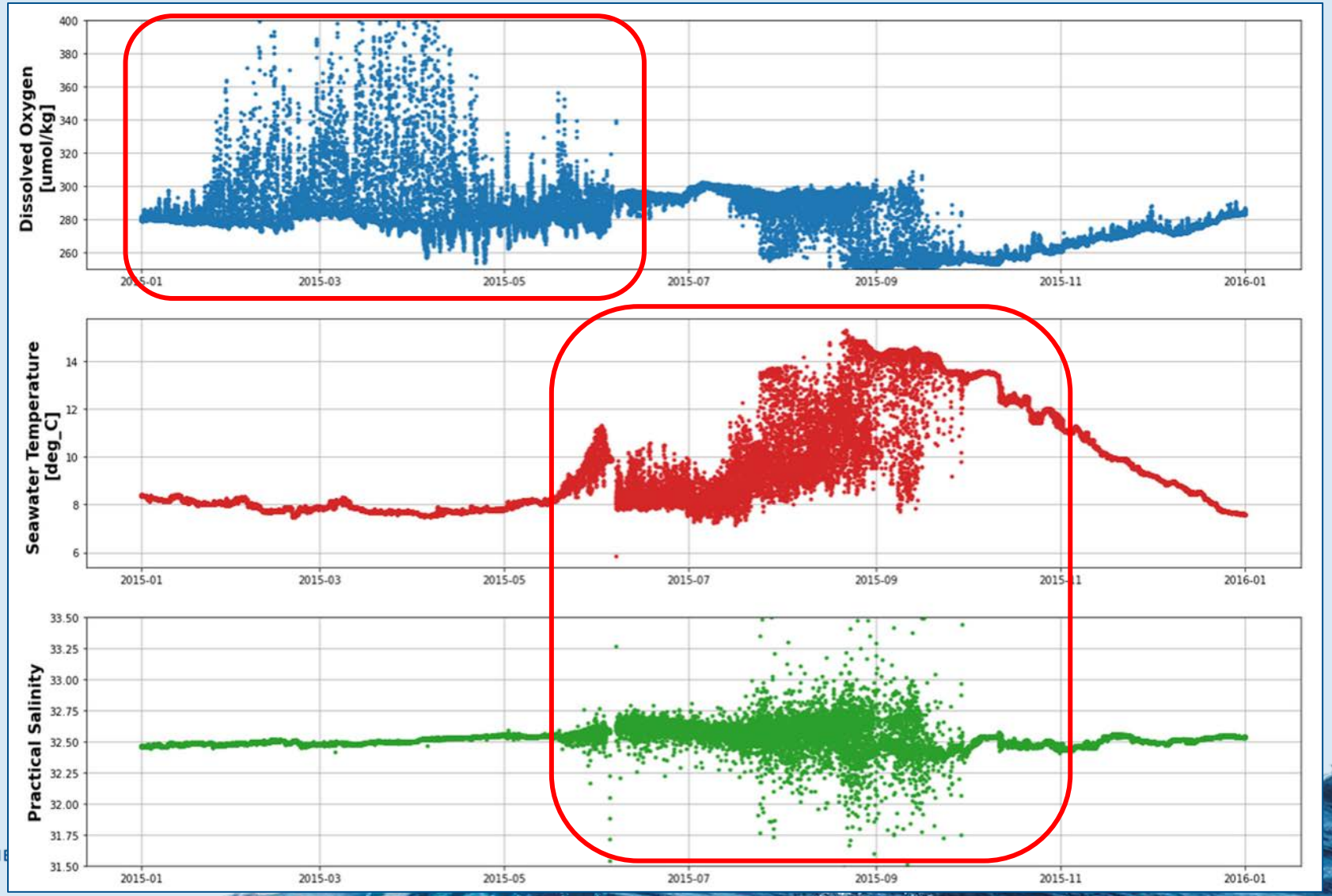
pressure	(time)	float64	nan nan nan nan ... nan nan nan nan
dissolved_oxyg...	(time)	uint8	29 29 29 29 29 ... 29 29 29 29 29
dissolved_oxyg...	(time)	uint8	29 29 29 29 29 ... 29 29 29 29 29

- Attributes: (70)

OCEANOBSERVATORIES.ORG

OCEAN OBSERVATORIES INITIATIVE

The Data



Annotations

Human-in-the-loop Assessment

- Added to the datasets for specified time periods by the operators
- Provide info on data quality, performance, issues, and context
- Available from the OOI Data Portal

Request

- Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12580/anno/find
- Parameters
  - refdes (required): reference designator
  - beginDT & endDT (optional): must be unix epoch time as microseconds
  - method (optional): only returns annotations for the given method for a given reference designator
  - stream (optional): only returns annotations for the given data stream for a given reference designator - stream

Annotations

- Example Request
  - Annotations for Ocean Station Papa Flanking Mooring A oxygen sensor from 2015-01-01 to 2016-01-01
  - Request url: https://ooinet.oceanobservatories.org/api/m2m/12580/anno/find?refdes=GP03FLMA-RIS01-02-DOSTAD000&beginDT=1420088401000&endDT=1451624401000

	@class	id	subsite	node	sensor	method	stream	beginDT	endDT	annotation	exclusionFlag	source	qcFlag	parameters
0	.AnnotationRecord	152	GP03FLMA	None	None	None	None	2015-06-06 22:40:00	2016-06-28 19:10:00	Deployment 3: Mooring was deployed 10-15m deep...	False	lgarzio@marine.rutgers.edu	0	[]
1	.AnnotationRecord	151	GP03FLMA	None	None	None	None	2014-06-17 06:00:00	2015-06-05 00:00:00	Deployment 2: Mooring was deployed 15-20m shal...	False	lgarzio@marine.rutgers.edu	0	[]
2	.AnnotationRecord	775	GP03FLMA	RIS01	03-DOSTAD000	None	None	2015-01-01 00:00:00	2015-06-05 00:00:00	Deployment 2: Dissolved oxygen noise increased...	False	lgarzio@marine.rutgers.edu	3	[]

```
2015-06-06 22:40:00 to 2016-06-28 19:10:00: Deployment 3: Mooring was deployed 10-15m deeper than planned.
2014-06-17 06:00:00 to 2015-06-05 00:00:00: Deployment 2: Mooring was deployed 15-20m shallower than planned. Upon recovery of platform, biofouling was apparent on shallow CTDs.
2015-01-01 00:00:00 to 2015-06-05 00:00:00: Deployment 2: Dissolved oxygen noise increased, possibly from biofouling.
```

- Answer to noisy data: **Biofouling!**

Metadata: Calibrations

Reference Designator based

- Since individual instruments are swapped during each mooring deployment & recovery, the calibration coefficients for a reference designator are different for each deployment. The way OOI operates is that it loads all the available calibration coefficients for a given reference designator.
- Recommend that limit an individual request to a single deployment - otherwise get a lot of calibrations
- Example request for Global Ocean Station Papa Flanking Mooring A oxygen optode for deployment 6
- Request url: https://ooinet.oceanobservatories.org/api/m2m/12587/asset/cal?refdes=GP03FLMA-RIS01-DOSTAD000&beginDT=2018-08-29T22:25:00.000Z&endDT=2018-08-30T22:54:00.000Z

	deploymentNumber	uid	calCoef	value	calFile
0	6	CGINS-DOSTAD-00228	CC_conc_coef	[-0.5159482, 0.9981074]	CGINS-DOSTAD-00228__20170918_Cal_Info.xlsx
1	6	CGINS-DOSTAD-00228	CC_csv	[0.00289589, 0.000120386, 2.34238e-06, 233.068...	CGINS-DOSTAD-00228__20170918_Cal_Info.xlsx

Metadata: Calibrations

Unique Identifier (UID) based

- Return a list for all of the calibrations available in the system for a particular instrument
- Request data for oxygen sensor with UID CGINS-DOSTAD-00228
- Request url: https://ooinet.oceanobservatories.org/api/m2m/12587/asset/cal?=CGINS-DOSTAD000-00228

	uid	calCoef	calDate	value	calFile
2	CGINS-DOSTAD-00228	CC_conc_coef	2013-08-04 00:00:00	[0.0, 1.0]	CGINS-DOSTAD-00228__20130804_Cal_Info.xlsx
4	CGINS-DOSTAD-00228	CC_csv	2013-08-04 00:00:00	[0.00313162, 0.000129918, 2.61329e-06, 233.397...	CGINS-DOSTAD-00228__20130804_Cal_Info.xlsx
0	CGINS-DOSTAD-00228	CC_conc_coef	2016-02-26 00:00:00	[0.0, 1.0]	CGINS-DOSTAD-00228__20160226_Cal_Info.xlsx
3	CGINS-DOSTAD-00228	CC_csv	2016-02-26 00:00:00	[0.00289589, 0.000120386, 2.34238e-06, 233.068...	CGINS-DOSTAD-00228__20160226_Cal_Info.xlsx
1	CGINS-DOSTAD-00228	CC_conc_coef	2017-09-18 00:00:00	[-0.5159482, 0.9981074]	CGINS-DOSTAD-00228__20170918_Cal_Info.xlsx
5	CGINS-DOSTAD-00228	CC_csv	2017-09-18 00:00:00	[0.00289589, 0.000120386, 2.34238e-06, 233.068...	CGINS-DOSTAD-00228__20170918_Cal_Info.xlsx

Metadata: Deployments

- Provide information on deployment numbers, times, cruises, and what instrument was deployed
- A deployment is defined as span of time a mooring or instrument were deployed and then recovered
- Reference Designator based
  - Request deployments for Global Ocean Station Papa Flanking Mooring A oxygen optode
  - Request url: https://ooinet.oceanobservatories.org/api/m2m/12587/events/deployment/inv/GP03FLMA/RIS01003-DOSTAD000
  - If no deployment number is given, returns a list of available deployments

```
API Endpoint: https://ooinet.oceanobservatories.org/api/m2m/12587/events/deployment/inv/GP03FLMA/RIS01/03-DOSTAD000

[1, 2, 3, 4, 5, 6, 7, 8, 9]
```

Metadata: Deployments

- Recommended to add in the deployment number
  - Request deployment 6 for Global Ocean Station Papa Flanking Mooring A oxygen optode
  - Request url: https://ooinet.oceanobservatories.org/api/m2m/12587/events/deployment/inv/GP03FLMA/RIS010 3-DOSTAD000
  - Can iterate through each deployment to build a table

	deploymentNumber	uid	assetId	latitude	longitude	depth	deployStart	deployEnd	deployCruise	recoverCruise
0	1	CGINS-DOSTAD-00129	2378	49.9795	-144.254	28.0	2013-07-21 22:44:00	2014-06-15 17:30:00	MV1309	MV1404
1	2	CGINS-DOSTAD-00127	2382	49.9775	-144.2463	28.0	2014-06-18 05:56:00	2015-06-05 14:53:00	MV1404	TN323
2	3	CGINS-DOSTAD-00394	3654	49.97667	-144.24617	28.0	2015-06-06 22:40:00	2016-06-28 15:02:00	TN323	RB1605
3	4	CGINS-DOSTAD-00130	2413	49.97434	-144.23972	30.0	2016-07-01 01:07:00	2017-07-18 15:13:00	RB1605	SR1710
4	5	CGINS-DOSTAD-00129	2378	50.02288	-144.36125	30.0	2017-07-13 23:00:00	2018-07-26 15:51:00	SR1710	SR1811
5	6	CGINS-DOSTAD-00228	2500	49.976145	-144.246335	30.0	2018-07-23 22:54:00	2019-09-29 15:54:00	SR1811	SKQ201920S
6	7	CGINS-DOSTAD-00379	3752	50.0227	-144.360783	30.0	2019-09-29 04:40:00	2021-07-26 18:24:00	SKQ201920S	SKQ202111S
7	8	CGINS-DOSTAD-00377	3756	49.977455	-144.244048	30.0	2021-07-23 19:43:00	None	SKQ202111S	None
8	9	CGINS-DOSTAD-00439	1208	49.9	-144.2	30.0	2022-05-11 00:00:00	None	SKQ202208S	None

Metadata: Vocab

OOI Vocabulary

- Provides more detailed information on the reference designator
- Includes the description names of the instrument at that reference designator, the array, the mooring, and node, as well as nomica
- Request vocab for Global Ocean Station Papa Flanking Mooring A oxygen
- Example url: https://ooinet.oceanobservatories.org/api/m2m/12586/vocab/inv/GP03FLMA/RIS01/03-DOSTAD000

	@class	vocabId	refdes	instrument	tocL1	tocL2	tocL3	manufacturer	model	mindepth	maxdepth
0	.VocabRecord	1292	GP03FLMA-RIS01-03-DOSTAD000	Dissolved Oxygen	Global Station Papa	Flanking Subsurface Mooring A	Mooring Riser	Aanderaa	Optode 4831	30.0	30.0

Metadata: Metadata

- The metadata contains such valuable information such as the available methods and streams (which are required to download the data), the particleKeys (the data variable names), parameter IDs, number of particles (data points), and the associated units
- Example request for Global Ocean Station Papa Flanking Mooring A oxygen
  - https://ooinet.oceanobservatories.org/api/m2m/12587/asset/cal?refdes=GP03FLMA-RIS01-DOSTAD000&beginDT=2018-08-29T22:25:00.000Z&endDT=2018-08-30T22:54:00.000Z

```
In [30]: metadata = get_metadata(refdes)
         metadata
```

Out[30]:

	pdId	particleKey	type	shape	units	fillValue	stream	unsigned	method	count
0	PD7	time	DOUBLE	SCALAR	seconds since 1900-01-01	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914
1	PD10	port_timestamp	DOUBLE	SCALAR	seconds since 1900-01-01	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914
2	PD11	driver_timestamp	DOUBLE	SCALAR	seconds since 1900-01-01	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914
3	PD12	internal_timestamp	DOUBLE	SCALAR	seconds since 1900-01-01	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914
4	PD14	dissolved_oxygen	FLOAT	FUNCTION	µmol kg-1	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914

- Results returns a table with 54 entries!

Preload: Metadata & Data Levels

- Data Levels refer to the level of processing that the given data parameter has undergone
  - Level 1, or L1 Data Products, are derived from L0 data, and provide data that has been calibrated using vendor-provided values or values derived from pre-deployment procedures, and that is in scientific units
  - Level 2, or L2 Data Products are derived quantities created via an algorithm that draws on multiple L1 Data Products. L2 data products may be based on data from the same or a combination of separate instruments.
- Preload
  - Contains detailed information on each parameter, queried based on the parameter ID
  - Get annotations for the GP03FLMA-RIS01-02-DOSTAD000 dissolved_oxygen (parameter ID = “14”)
  - Example request: **https://ooinet.oceanobservatories.org/api/m2m/12575/parameter/14**

Preload: Metadata & Data Levels

- Example request: **https://ooinet.oceanobservatories.org/api/m2m/12575/parameter/14**


```
{'name': 'dissolved_oxygen',
 'display_name': 'DO from Onboard Calculation - Corrected',
 'netcdf_name': 'dissolved_oxygen',
 'standard_name': 'moles_of_oxygen_per_unit_mass_in_sea_water',
 'description': 'Dissolved Oxygen Concentration from the Stable Response Dissolved Oxygen Instrument is a measure of
the concentration of gaseous oxygen mixed in seawater. This data product is corrected for salinity, temperature, and
depth.',
 'id': 14,
 'data_product_identifier': 'DOXYGEN_L2',
 'precision': 4,
 'fill_value': {'value': '-9999999'},
 'unit': {'value': 'µmol kg-1'},
 'data_level': 2,
 'code_set': None,
 'value_encoding': {'value': 'float32'},
 'parameter_type': {'value': 'function'},
 'parameter_function': {'id': 61,
  'name': 'do2_salinity_correction',
  'function_type': {'value': 'PythonFunction'},
  'function': 'do2_salinity_correction',
  'owner': 'ion_functions.data.do2_functions',
  'description': 'Salinity and pressure correction to dissolved oxygen producing L2 DOCONCS',
  'qc_flag': None},
 'data_product_type': {'value': 'Science Data'},
 'dimensions': [],
 'parameter_function_map': '{"DO": "dpi_DOCONCS_L1", "SP": ["dpi_PRACSAL_L2", "dpi_SALSURF_L2"], "lon": "CC_lon", "
P": ["PD2606", "dpi_PRESWAT_L1", "PD17"], "T": ["dpi_TEMPWAT_L1", "dpi_TEMPSRF_L1"], "lat": "CC_lat"}',
 'visible': True}
```

Metadata & Preload

- With the Data Levels queried from **Preload** for each of the parameters returned with the **Metadata**, we can filter for the different Data Levels. Below we filtered the metadata for the GP03FLMA-RIS01-03-DOSTAD000 for L1 & L2 data products
  - This allows us to cut down on the size of the data request and resulting data files

	pdId	particleKey	type	shape	units	fillValue	stream	unsigned	method	count	
4	PD14	dissolved_oxygen	FLOAT	FUNCTION	µmol kg-1	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914	2014-06
7	PD940	estimated_oxygen_concentration	FLOAT	SCALAR	µmol L-1	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914	2014-06
18	PD2843	dosta_abcdjm_cspp_tc_oxygen	FLOAT	FUNCTION	µmol L-1	-9999999	dosta_abcdjm_sio_instrument	False	telemetered	3914	2014-06
23	PD14	dissolved_oxygen	FLOAT	FUNCTION	µmol kg-1	-9999999	dosta_abcdjm_sio_instrument_recovered	False	recovered_host	152192	2013-07
26	PD940	estimated_oxygen_concentration	FLOAT	SCALAR	µmol L-1	-9999999	dosta_abcdjm_sio_instrument_recovered	False	recovered_host	152192	2013-07
37	PD2843	dosta_abcdjm_cspp_tc_oxygen	FLOAT	FUNCTION	µmol L-1	-9999999	dosta_abcdjm_sio_instrument_recovered	False	recovered_host	152192	2013-07

- **Important:** Notice that **time** was filtered out. This is because they don't have a defined data product level. An additional wrinkle is that **time** is NOT the default dimension of delivered netCDF files - this means it needs to be specifically requested for data requests. This is something to be aware of when requesting only specific data variables.

Further Information

- Presentation available on gitHub at **github.com/reedan88/OOI-NE-Pacific-Community-Workshop**
  - Includes python package for simplified interaction with the OOI M2M & example notebook
  - Also includes jupyter scripts used to create “Global Ocean Station Papa: Assets & Datasets” poster

- OOI Data Explorations
  - Available on the OOI github at **github.com/oceanobservatories/ooi-data-explorations**
  - Includes code to both request and process data from OOI as well as some QC routines
  - Code in both python, MatLab, and R

- Supplemental & Complementary Datasets
  - Discrete water sampling data available on the OOI Alfresco Web Document Server
  - alfresco.oceanobservatories.org: OOI > Array > Cruise Data > Cruise > Ship Data > Water Sampling

- Other Sources for OOI Data
  - OOI Data Explorer website & ERDDAP server - see “**Data Explorer”** Breakout Session
  - Raw Data Server, Camera Data, & Hydrophone Data - see “**Raw Data Server”** Breakout Session

OCEAN
OBSERVATORIES
INITIATIVE

Questions?