OOI data types NOT on Data Explorer:

- HD video (CAMHD)
- Bioacoustic sonar (ZPLSC/G)
- Still imagery (CAMDS)
- Hydrophones (HYDBB/LF)
- Seismometers (OBSBB/OBSSP)
- Fluid and particulate samples
- Cruise Data
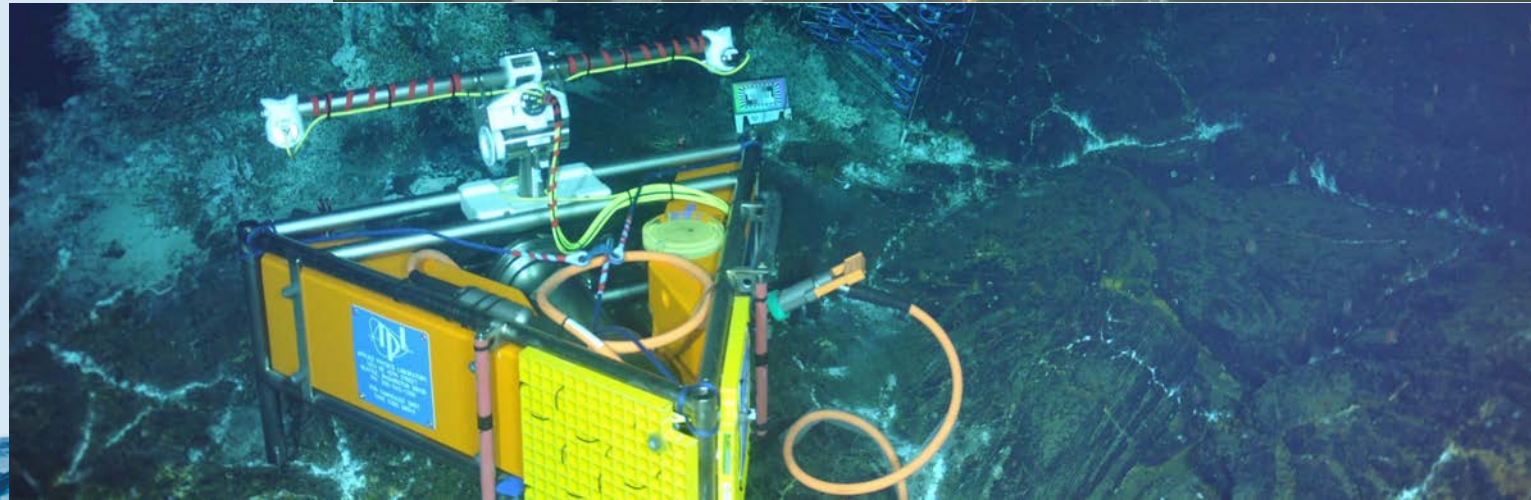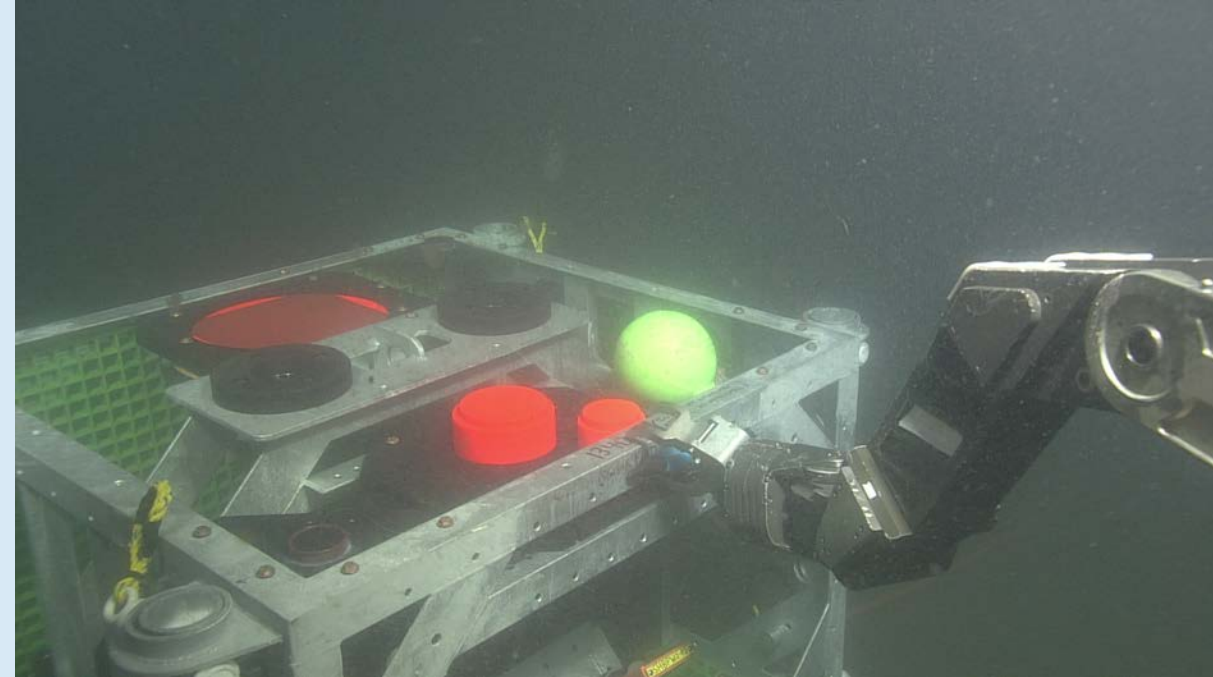- PI instruments and community data products

OOI OCEAN OBSERVATORIES INITIATIVE

How do I find an instrument?

- Unless you already know the route to the specific platform and instrument you're looking for, you should begin on the OOI website or data portal

- Start on the landing page of the OOI Data Portal or Data Explorer and navigate to the platform, instrument, or data that you are interested in

- Note the "Reference Designator" so that you can see the full name of the instrument and node number (e.g., *RS03ASHS-PN03B-06-CAMHDA301*)

- Use that as a reference when navigating within the raw data archive

OCEAN OBSERVATORIES INITIATIVE

What is the Raw Data Archive?

- Data storage and access for files that cannot be delivered via Data Explorer

- This server provides access to the raw data to:
    - Allow SMEs to assist with quality control of specialized instrumentation
    - Allow users to perform their own analyses using their own scripts or software
    - Allow MIOs to confirm operation of deployed instrumentation

- "Raw" indicates data as they are received directly from the instrument, in instrument-specific format
    - May contain data for multiple sensors (interleaved), be in native sensor units (e.g., counts, volts) or have processing steps already performed within the instrument (e.g., primary calibration)
    - This also includes large format audio/visual data (e.g., HD video, still images, and hydrophone data)
    - Data may not be in familiar, vendor formats (due to data delivery method or OOI processing)

OCEAN OBSERVATORIES INITIATIVE

NSF

What is found on the Raw Data Archive?

- All raw data from all platforms and instruments are stored by OOI, including Antelope data (hydrophones and seismometers). The OOI also delivers a subset of the total stored raw data via the Raw Data Archive.

- Data in the Raw Data Archive (based on the current size) are:

  - All uncabled raw data for an initial period of 10 years

  - All cabled raw data (minus Antelope and HD Video) for an initial period of 10 years

  - An initial period of 6 months of broadband hydrophone (HYDBB) data

  - An initial period of 6 months of full-resolution HD Video data (.mov files)

  - An initial period of 10 years of compressed HD Video data (.mp4 files)

- Data removed from the Raw Data Archive for space reasons are <u>not ever deleted</u>. The data not directly accessible in the Raw Data Archive will be made available to users upon request

OCEAN OBSERVATORIES INITIATIVE

NSF

What is NOT on the Raw Data Archive?

- OOI low-frequency hydrophones (HYDLF) and seismometer data (OBSBB, OBSSP) are currently only hosted by IRIS (Incorporated Research Institutions for Seismology: http://ds.iris.edu/ds/).

- Shipboard data collected during OOI deployment and recovery cruises are stored on the Alfresco document management system (Cruise Data Archive)

- When a user submits a Help Desk request for raw data that are not available on the NAS, the OOI cyberinfrastructure team will provide an alternate data delivery method, which will depend on the size of the request

OCEAN OBSERVATORIES INITIATIVE

NSF

How is the Raw Data Archive organized?

- The archive is a mirror of the data repository where all raw data enters the system

- Uncabled mooring data are organized by deployment/recovery number for each uncabled platform (e.g., D00001/R00001), with subfolders for each node (control computer), containing another subfolder for each attached instrument which contains all data for that instrument from that deployment

- Uncabled mobile assets (i.e., gliders and AUVs) are organized similarly, but the subfolders below the deployment/recovery number are organized based on the glider's internal science computer file structure:

  - Telemetered data folders are titled D0000*, and contain subfolders called archive, from-glider, logs, and merged-from-glider.

  - Recovered data folders (R0000*) are titled cache, merged, dvl, flight, and science.

  - "Merged" folders contain most of the science instrument and engineering data of interest, although the glider ADCP data is contained in the "dvl" subfolder.

- Cabled data are pulled from two archives and merged. Most cabled data are organized by site and node, which refers to the alphanumeric ID of the junction box attached directly to the undersea cable that aggregates, time-stamps, and routes the data from all instruments connected to that node. Instrument data are then organized by date (yyyy/mm/dd)

OCEAN OBSERVATORIES INITIATIVE

How do I access metadata?

- OOI platform and instrument metadata are currently provided when users request processed data, or via the asset management page of the OOI Data Portal.

- Vendor-provided calibration or instrument setting sheets that may be required for data analysis can be found via Alfresco.

- Calibration values can be found on GitHub (https://github.com/ooi-integration/asset-management/tree/master/deployment), and copies of the sheets that accompany new or refurbished instruments can be provided upon request by contacting the HelpDesk.

- Any additional required metadata can be requested from the program via Help Desk or the Discourse forum.

How do I grab all data from a directory?

- In order to recursively download an entire raw data directory from rawdata.oceanobservatories.org, you can use the following script:

```
wget -r -np -e robots=off URL
```

- Notes:

  - -r signifies that wget should recursively download data in any subdirectories it finds.

  - -np prevents files from parent directories from being downloaded.

  - -e robots=off tells wget to ignore the robots.txt file. If this command is left out, the robots.txt file tells wget that it does not like webcrawlers and this will prevent wget from working.

- Here is a complete example:

```
wget -r -np -e robots=off https://rawdata.oceanobservatories.org/files/CE01ISSP/D00001/
```

- While this is a simple example, there are many more options you can try when using wget

OCEAN OBSERVATORIES INITIATIVE

How do I use the raw data files?

- Depends on the instrument/file type

- See website FAQ or forum posts for instructions

- Some file types require specific software, others have code available to open files locally

OCEAN OBSERVATORIES INITIATIVE

NSF

HD Video (CAMHD)

- Cabled high-definition video is available in two formats:

  - Uncompressed full-HD video files are in .mov file format, which are very large and may take a long time to download

  - Smaller files created using lossless compression are in .mp4 format, which are still large but are more easily downloaded and played

- Log files with the pan & tilt routine and other metadata are in an associated subfolder

2022-05-27 21:15:00

OCEAN OBSERVATORIES INITIATIVE

NSF

Still images (CAMDS)

- Since 2019, all of the cabled units (instrument code CAMDS-B and CAMDS-C) are SubC Rayfins with Aquorea LED lamps and ROS PT-25 pan & tilt units

- Uncabled cameras were Kongsberg units and are now a mix of models

- Image files are available on the Raw Data server in .png format, organized by reference designator and YYYY/MM/DD subfolders

OCEAN OBSERVATORIES INITIATIVE

NSF
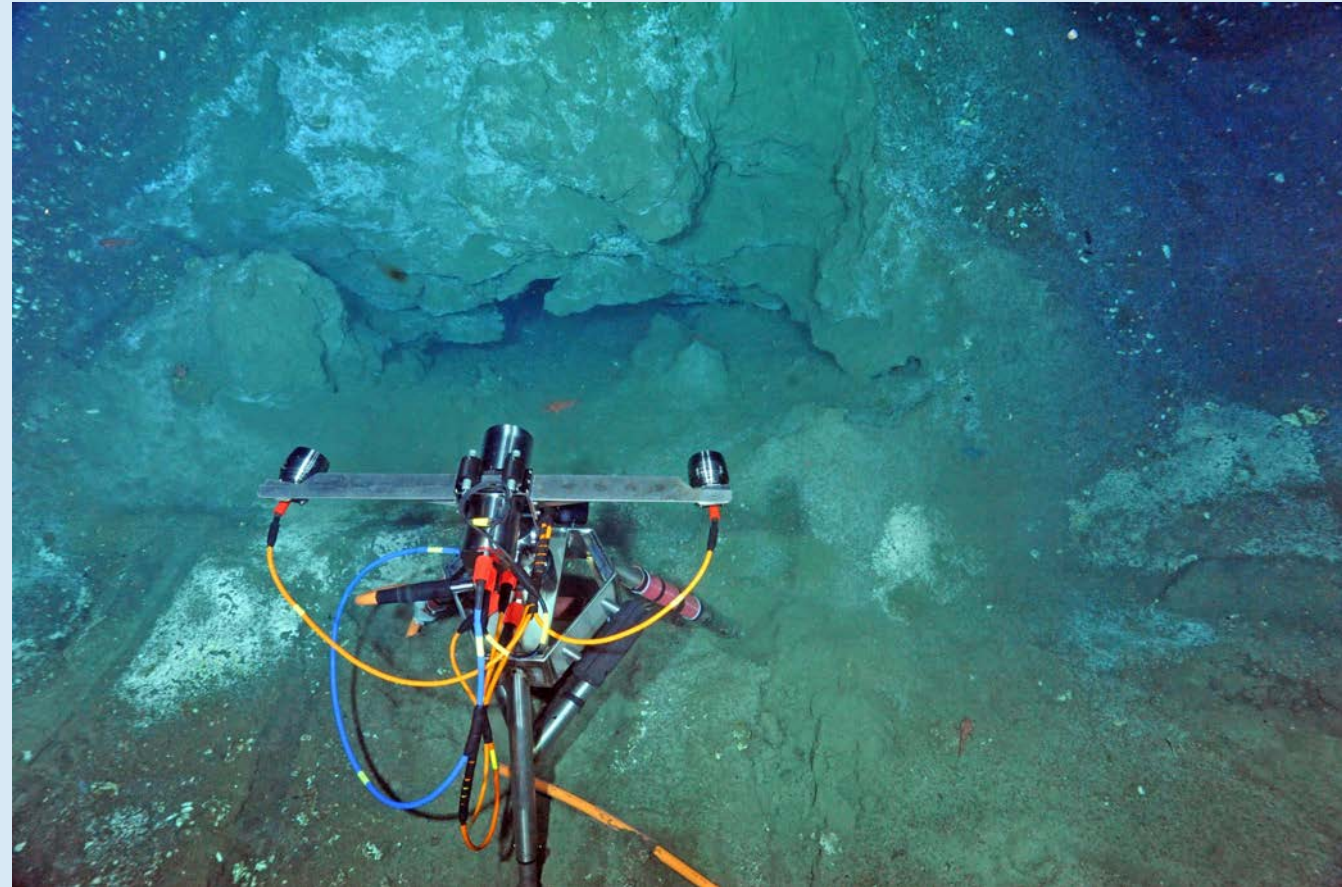
IRIS (seismic data)

- IRIS is a consortium of universities dedicated to the acquisition, management, and distribution of seismological data

- All OOI seismic-related and low-frequency (LF) hydrophone data collected by the Regional Cabled Array are deposited in IRIS

- They are also importing bottom pressure and tilt unit (BOTPT aka tiltmeter) data

- Many types of tools and links to access and use the seismic data

contact us | sign in

IRIS Incorporated Research Institutions for Seismology

DATA   INSTRUMENTATION   EDUCATION   ABOUT

SAGE

DATA SERVICES / NODES / DMC / TOOLS

OOI OCEAN OBSERVATORIES INITIATIVE

Sonar Data

- Cabled bioacoustic sonar data (ZPLSC-B; modified Kongsberg EK-60 echosounders) are in vendor format .raw files

- Raw files can be opened using EchoView or vendor software, or open-source Echopype software

- A media server is in development that will use Echopype to create echograms and allow users to browse the sonar data more easily

echopype

Endurance Oregon Offshore Surface Mooring

Oregon Offshore Cabled Shallow Profiler Mooring Bio-acoustic Sonar: 200kHz
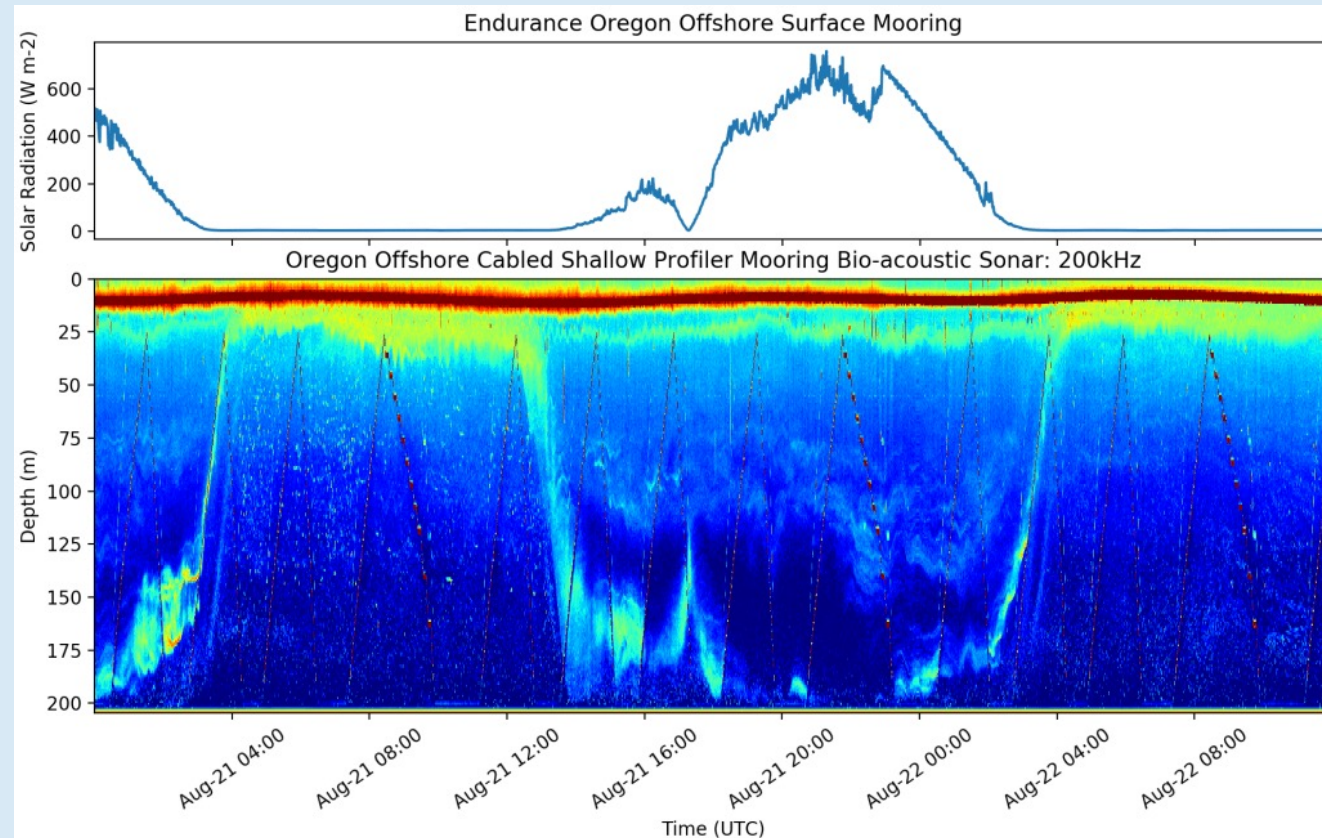
Coastal and Global bioacoustic data

- Global Station Papa has Bio-acoustic Sonar instruments (ZPLSG) to sense zooplankton in the upper ~300 meters from the Apex Profiler Mooring

  - These data are only available from the Raw Data Archive, including processed NetCDF and echogram image files

  - In the future, the NetCDF files will also be available through OOINet and Data Explorer)

- To access these data, as well as data quality annotations:

  - Navigate to the Papa Apex Profiler Mooring in the Raw Data Repo: https://rawdata.oceanobservatories.org/files/GP02HYPM/

  - Go into the Recovery "R" folder for the deployment that you are interested in, then go into the instruments subfolder (e.g., deployment 7): https://rawdata.oceanobservatories.org/files/GP02HYPM/R00007/instruments/

  - You will see two ZPLSG instruments (upward- and downward-looking from 150-m depth), identified by serial number. Within each ZPLSG instrument folder, there are subfolders DATA (for the raw data) and processed, and if there were any notable annotations, you'll see a ReadMe file (e.g., https://rawdata.oceanobservatories.org/files/GP02HYPM/R00007/instruments/ZPLSG_sn55107/)

  - Within the processed folder there are subfolders organized by week with start and end dates. Within these subfolders are hourly-averaged (YYMMDDHH) and daily NetCDF files, with an echogram .png image for the week: https://rawdata.oceanobservatories.org/files/GP02HYPM/R00007/instruments/ZPLSG_sn55107/processed/20210103-20210110/GP02HYPM_LOWER_Bioacoustic_Echogram_20210103-20210110_Calibrated_Sv.png

Hydrophones

- There are six broadband hydrophones (HYDBB) on the Regional Cabled Array

- Broadband hydrophone files are currently in .mseed format, in 5-minute increments

- Tools for opening those files can be found on the IRIS site (http://ds.iris.edu/ds), among other locations

- An audio player/visualization tool is in development that will allow the files to be viewed, heard, and then downloaded in FLAC or .wav format, which should be able to be opened and listened to using most audio player software

OCEAN OBSERVATORIES INITIATIVE

Fluid and Particulate Samples

- Datasets contain processed analytical data and metadata for a subset of Cabled Array instruments that collect fluid or particulate samples that are brought back to shore and analyzed in a laboratory.

- Numerical and other processed analytical results are generally provided in spreadsheet form.

- Associated metadata files including analytical protocols, sampling and recovery notes, and sample log files are presented in .pdf format.

- As-deployed images of instruments are provided in .jpg format, with image metadata summarized in spreadsheet-based catalogs.

- https://oceanobservatories.org/core-instrument-analytical-results/

OCEANOBSERVATORIES.ORG

OCEAN OBSERVATORIES INITIATIVE

Fluid and Particulate Samples

- Osmosis-based Water Sampler (OSMOI):

    - Long-term sampling of diffuse and/or pore fluids at the seabed by drawing those fluids into small-bore, capillary tubing

    - Deployed at Southern Hydrate Ridge to sample pore fluids associated with methane seeps, and at the ASHES vent field on Axial Seamount collocated with the 3-D Temperature Array (TMPSF) to sample diffuse fluids associated with hydrothermal flow

- Hydrothermal Vent Fluid Interactive Sampler (RASFL):

    - Collects fluid samples from hydrothermal vents at the International District Vent Field on Axial Seamount. It is collocated with the Particulate DNA Sampler and has an onboard temperature sensor (D1000) which continuously measures vent fluid temperatures.

- Particulate DNA Sampler (PPSDN):

    - Filters and preserves particulate matter from venting hydrothermal fluids at the International District Vent Field on Axial Seamount

    - Genetic analysis of the particulates provides information on the microbial communities associated with hydrothermal vents

- Benthic Fluid Flow (FLOBN-C and FLOBN-M):

    - Measures fluid flow at the sediment-water interface by monitoring the dilution of a chemical tracer injected into fluids at the seabed and collected in small-bore, capillary tubing by osmotic pumps

    - The "C" series of this instrument is a Chemical and Aqueous Transport (CAT) meter, and the "M" series of this instrument is a Multiple Orifice Sampler and Quantitative Injection Tracer Observer (MOSQUITO)

    - Both are deployed on or near methane seeps at the Southern Hydrate Ridge seabed

    - Recovered fluids are analyzed for chemical tracer and major ion concentrations, and flow rates are calculated

OCEAN OBSERVATORIES INITIATIVE
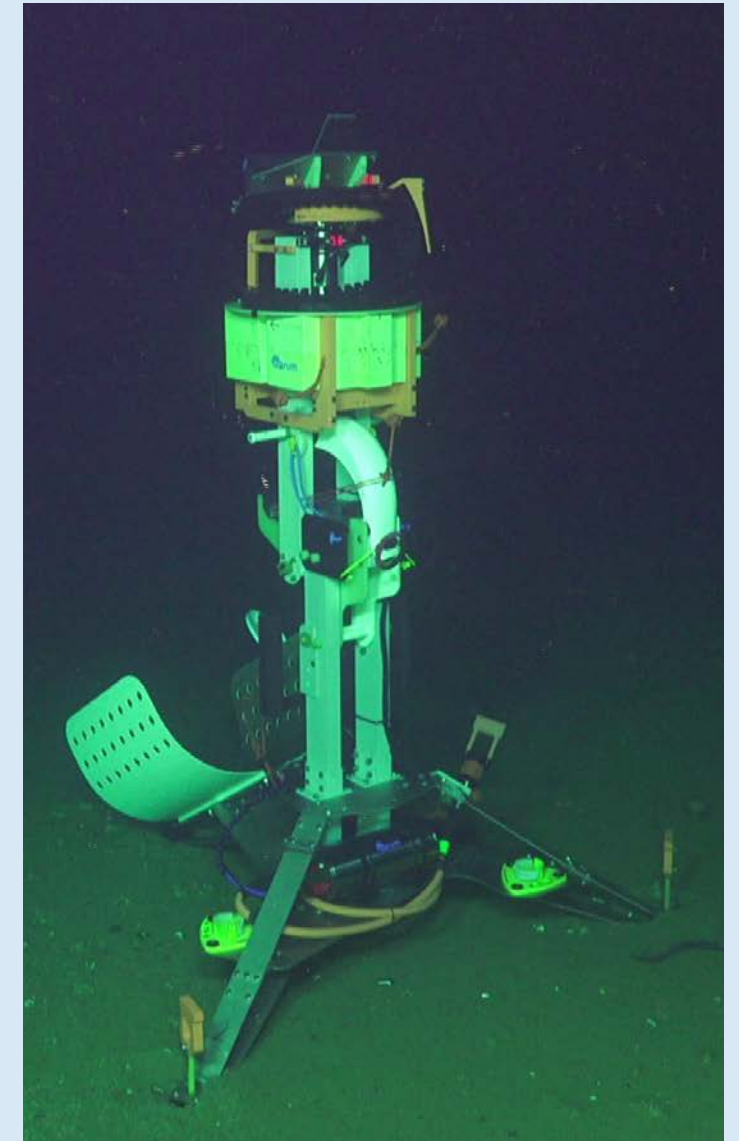
NSF

Cruise data (discrete water samples)

- OOI also collects CTD and discrete water sampling data during turn-around cruises

- All cruise data and array documentation can be accessed and browsed directly at: https://alfresco.oceanobservatories.org/

- This repository includes:

  - Vendor documentation on delivery, recalibration, and refurbishment of instruments,

  - OOI internal design documents

  - Pre- and post-deployment testing documentation for all instruments and platforms

  - Cruise plans and reports for all turn-around cruises to service the arrays

  - All data collected on those cruises, including CTD cast and discrete bottle sample data (which are found in Discrete Summary CSV files in "Water Sampling" subfolders (e.g., Cabled Array **>** Cruise Data **>** Cabled-12_TN393_2021-07-30 **>** Ship Data **>** Water Sampling **>** Cabled-12_TN393_Discrete_Summary.csv)

OCEAN OBSERVATORIES INITIATIVE
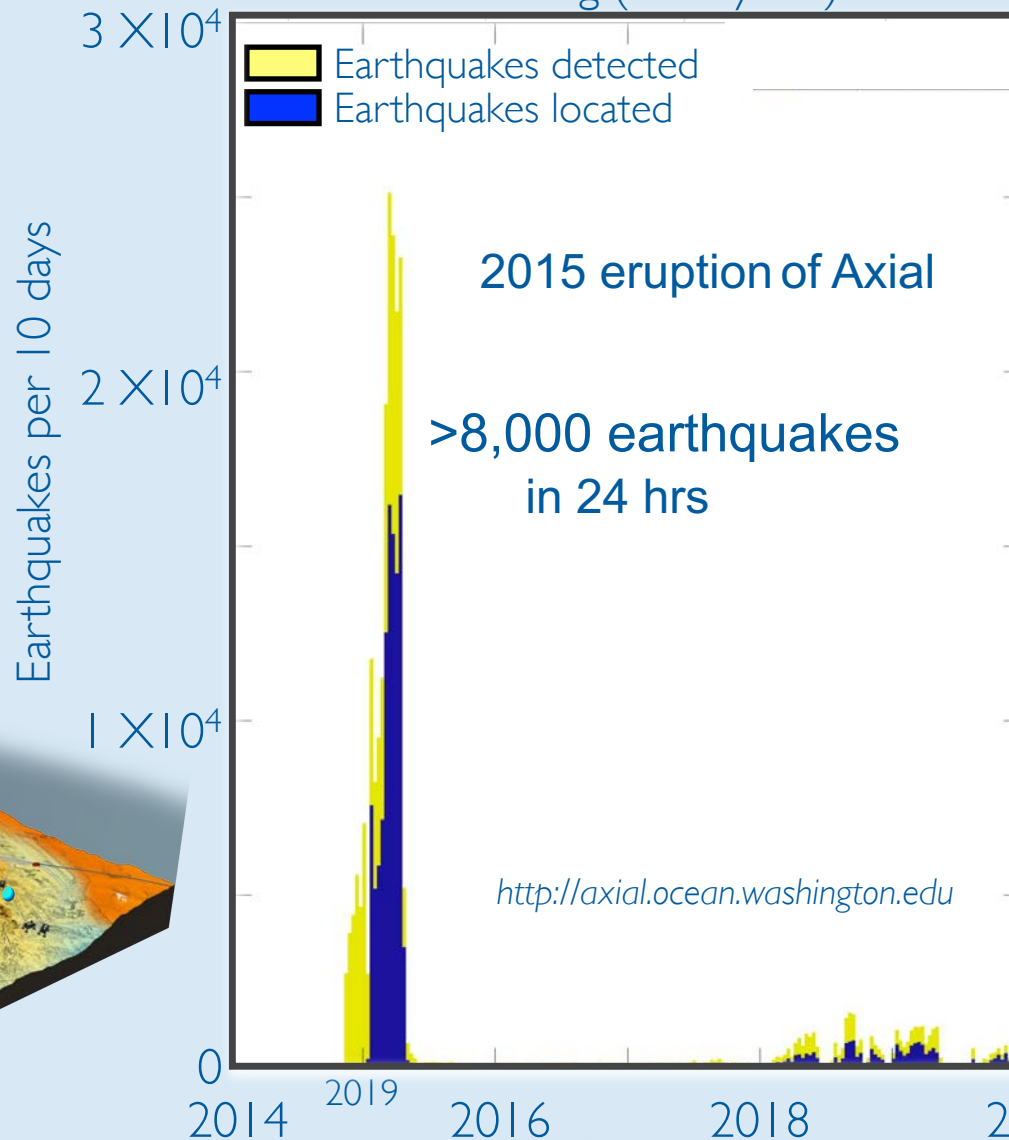
NSF

PI Instruments

MARUM Overview Sonar
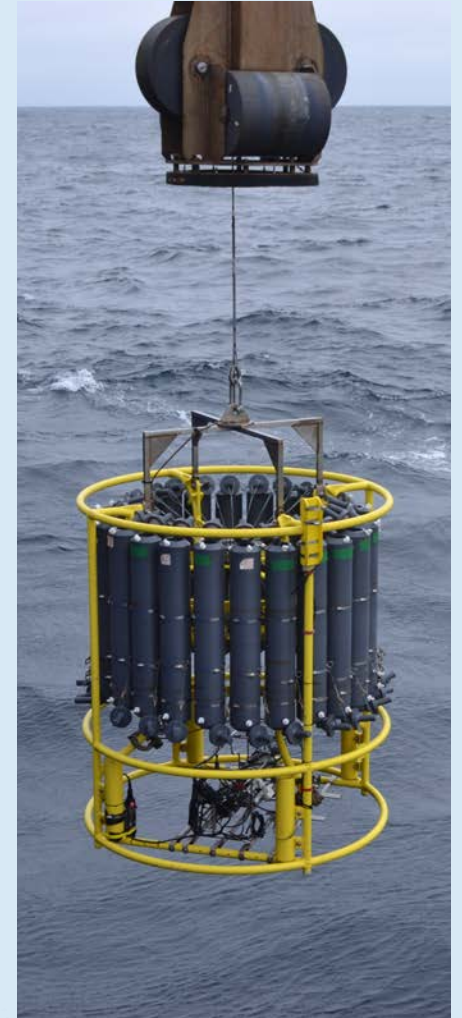
- Datasets compiled by Principal Investigators who have added instruments onto OOI arrays

- The National Science Foundation provides funding through its core programs to encourage scientists to add new instruments or platforms onto the OOI Observatory
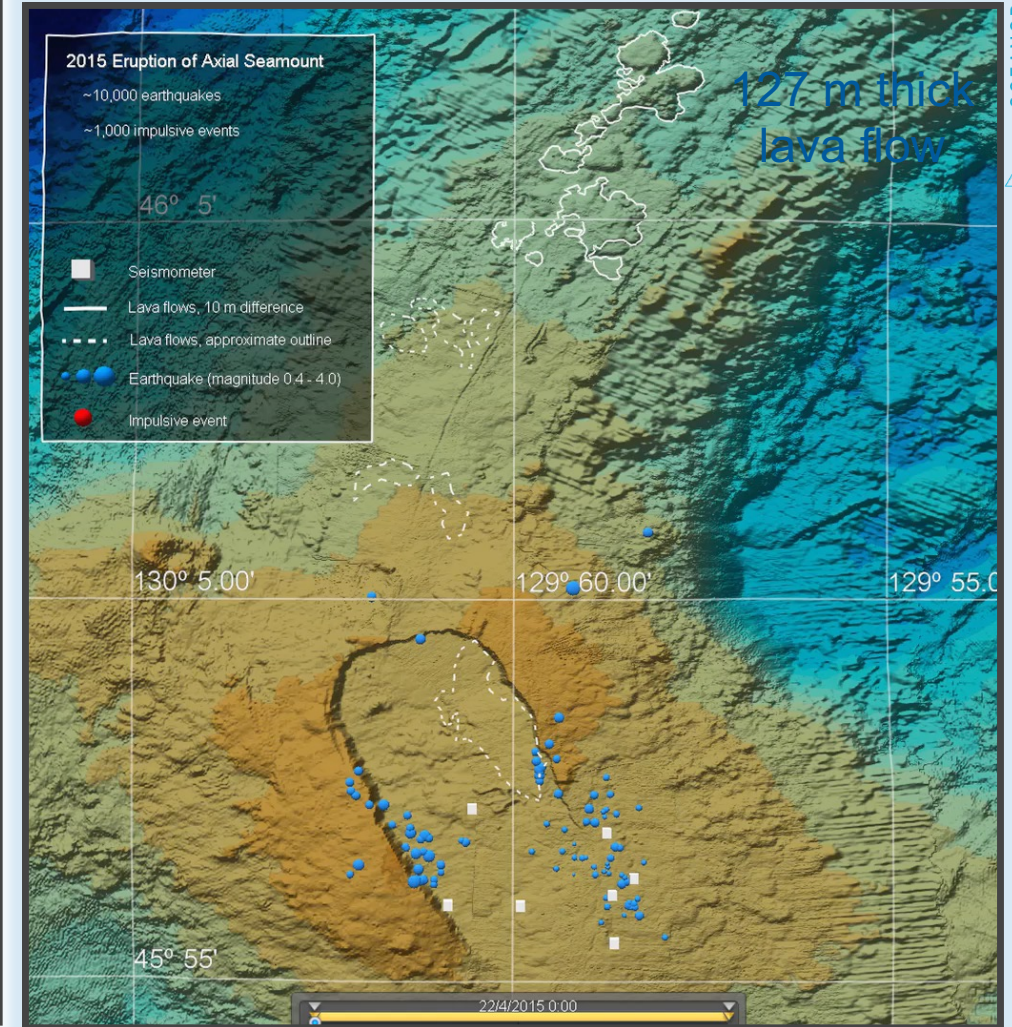
- https://oceanobservatories.org/instruments/

- https://oceanobservatories.org/pi-added-instruments/

OCEAN OBSERVATORIES INITIATIVE

NSF

Community Products

Full Catalog (10 day bin)

Wilcock et al., Science *2016*

OCEANOBSERVATORIES.ORG

Earthquakes per 10 days

- Earthquakes detected
- Earthquakes located

2015 eruption of Axial

>8,000 earthquakes
in 24 hrs

http://axial.ocean.washington.edu

3 X10⁴
2 X10⁴
1 X10⁴
0

2014   2019   2016   2018   2020

2015 Eruption of Axial Seamount
~10,000 earthquakes
~1,000 impulsive events

46° 5'

- ☐ Seismometer
- ─ Lava flows, 10 m difference
- ┈ Lava flows, approximate outline
- • Earthquake (magnitude 0.4 - 4.0)
- ● Impulsive event

127 m thick
lava flow

130° 5.00'   129° 60.00'   129° 55.0

45° 55'

22/4/2015 0.00

● Earthquakes   ● Hydroacoustic events (explosions)

OCEAN OBSERVATORIES INITIATIVE

NSF

Questions?