Managing Metadata and Data for the OOI

Grace Agnew, Associate University Librarian for Digital Library Systems (gagnew@rutgers.edu)

Chad Mills, Digital Library Architect (cmmills@rulmail.rutgers.edu)

RUTGERS
University Libraries

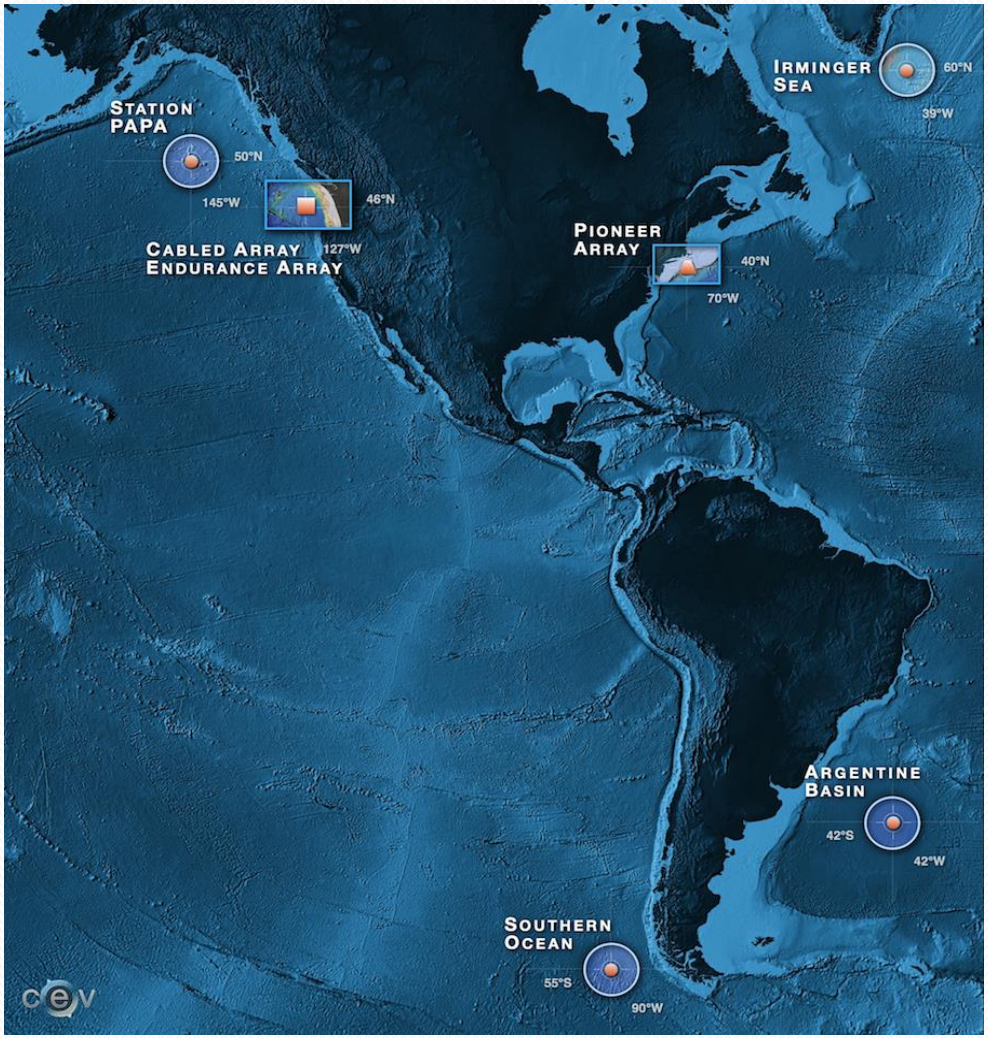
Introduction

- What is Ocean Observatories Initiatives (OOI)?
- OOI Project Team Members:
  - Consortium for Ocean Leadership
  - Woods Hole Oceanographic Institution
  - Oregon State University
  - University of Washington
  - Rutgers University
- Research Fields: (Global, Regional, Coastal and Cyberinfrastructure)
- Data Definition (Raw data, L0, L1 and L2)

oceanobservatories.org

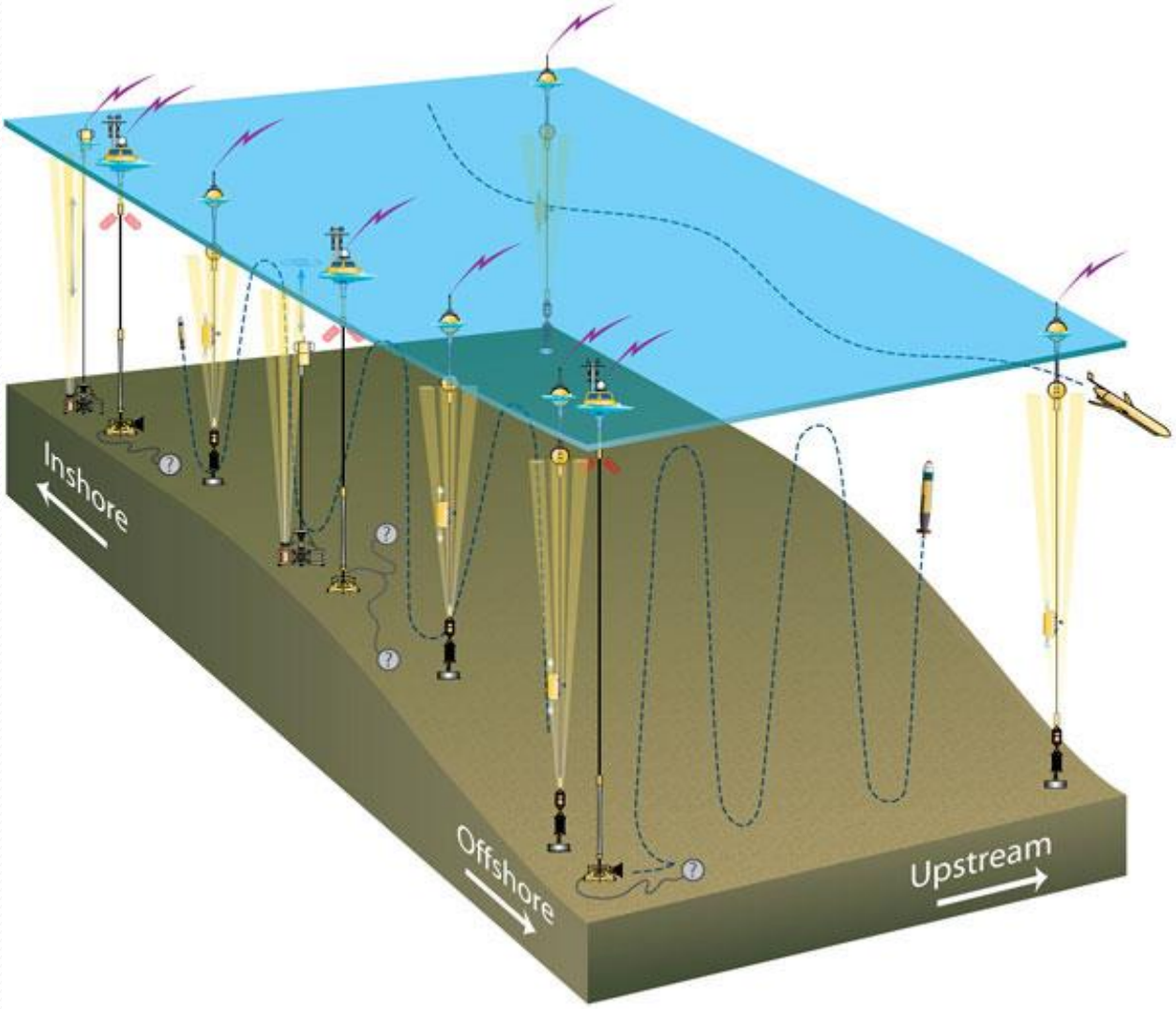
Observatory Footprint

- 7 Arrays –Atlantic and Pacific oceans
- 75 Platforms/Sub-Sites
- 800+ instruments
- 10,000+ data products initially



Credit: OOI Cabled Array program and the Center for Environmental Visualization, University of Washington

Pioneer Coastal Array



List of Data Products

- OOI collects 108 types of data products (temperature, salinity, etc.)

- Data products are broken up in three primary sampling regimes
  - Air-Sea Interface
    - Air Temperature; Rain Rate; Sea Surface Conductivity; etc.
  - Seafloor/Crust
    - Hydrogen Concentration; HD Video; Seafloor Pressure; etc.
  - Water Column
    - Conductivity; Echo Intensity; Signal Absorption; etc.

Data Products

Data Types

- RAW data; unprocessed
- Level 0/L0 data; processed
- Level 1/L1 data; edited
- Level 2/L2 data; derived

Once an instrument is deployed, that instrument will send new data to the OOI cyberinfrastructure every 15 minutes. This data is referred to as RAW or unprocessed data.

RAW data is used by software parsers to remove unwanted information that was recorded. The resulting data is referred to as Level 0/L0 or processed data.

The OOI is expected to record and maintain multiple petabytes of RAW and L0 data.

Libraries' Involvement

Citable, referenceable links to persistent data products.

DOI and metadata were the tools selected for this project

http://dx.doi.org/doi:10.7282/T3C53NKV

The libraries mint DOIs for RUcore resources. Also, as part of the RUcore's DOI implementation, resources are indexed and discoverable through many services, for example, Google, Google Scholar, Thompson Reuters, DataCite, etc.

Initial involvement – Minting DOIs—led to many other things, including dynamic but durable metadata for research data.

RUcore Implementation

Rucore – Fedora-based repository framework with APIs for creation, discovery and access to resources.

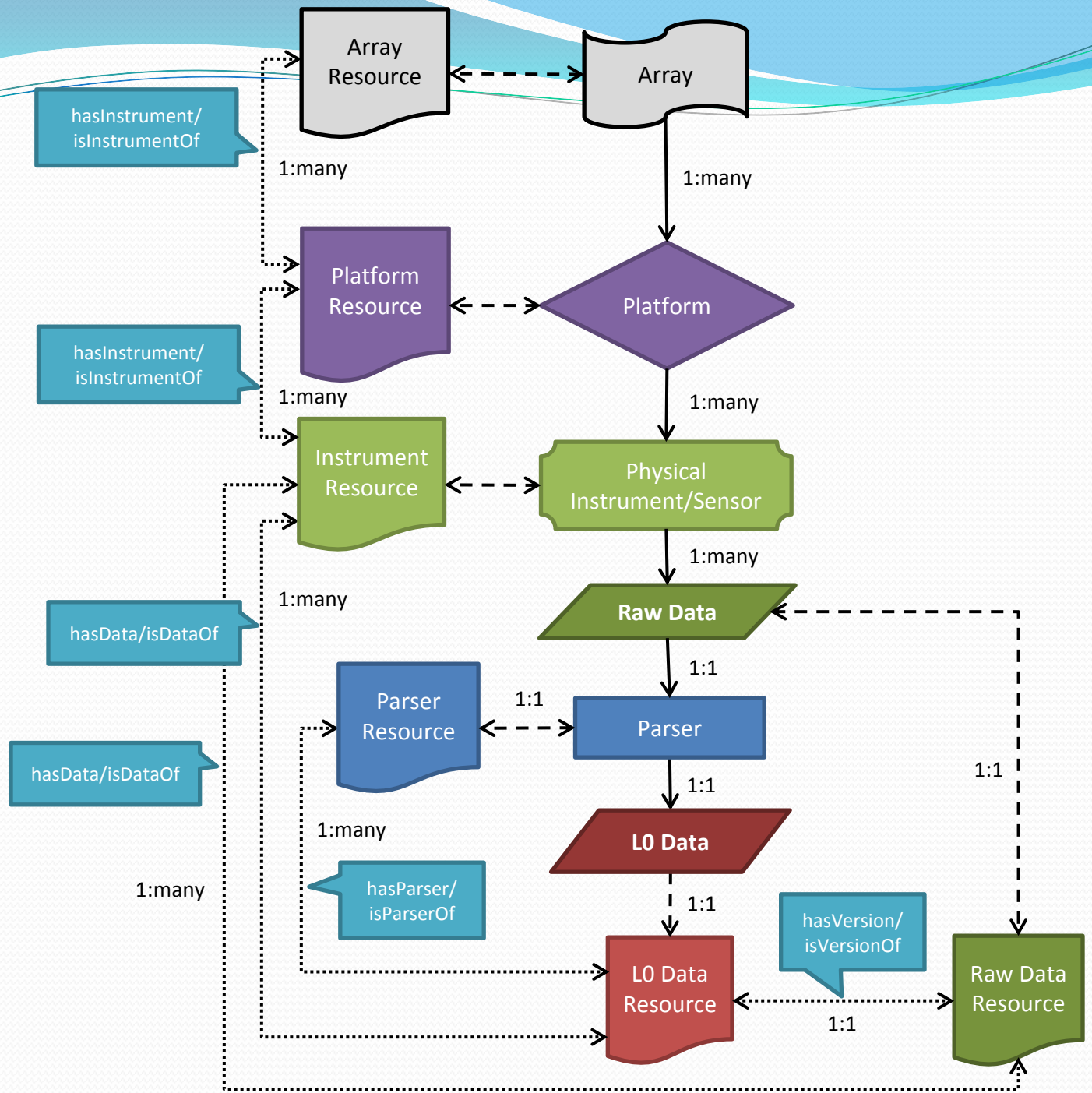
Building Block approach for services that are reusable, customizable for different projects.

- **Developing a research data model**
- **Developing a metadata application profile**
- **API for the OOI infrastructure to automatically create resources in RUcore for data products**

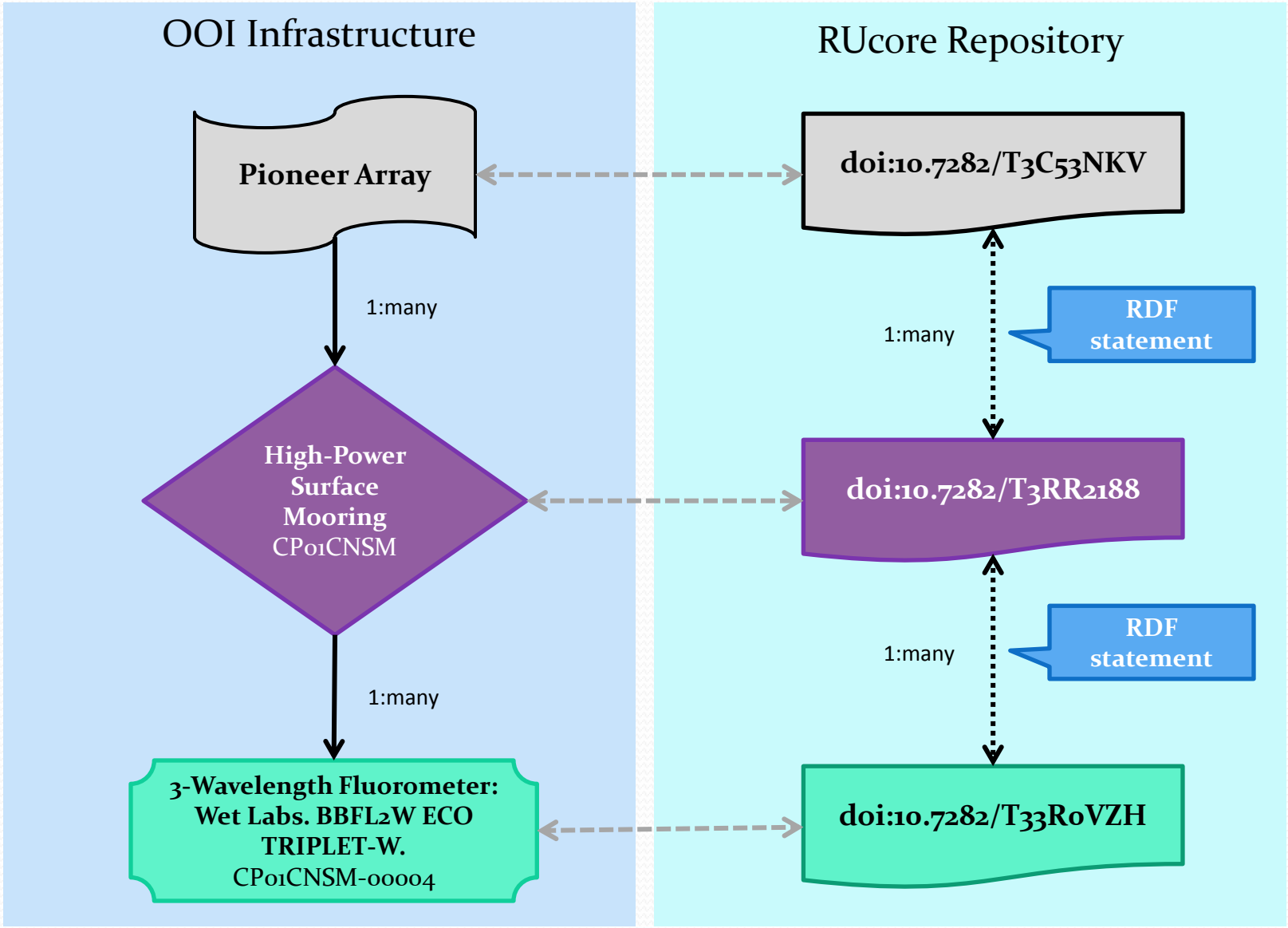
Creating a Research Data Model

- Documents the different types of content
- Describes relationships and dependencies between content
- Needed to define, describe, disseminate and discuss content

- **Why is it important?**
  - Clarifies requirements
  - Rationalizes design – shared understanding
  - The data defines the metadata, not the other way about


Array Resource
Array
hasInstrument/ isInstrumentOf
1:many
1:many
Platform Resource
Platform
hasInstrument/ isInstrumentOf
1:many
1:many
Instrument Resource
Physical Instrument/Sensor
1:many
1:many
Raw Data
hasData/isDataOf
1:1
Parser Resource
1:1
Parser
hasData/isDataOf
1:1
1:many
L0 Data
1:many
hasParser/ isParserOf
1:1
hasVersion/ isVersionOf
L0 Data Resource
Raw Data Resource
1:1

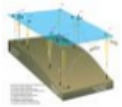
Data Model in Action



Metadata Design

- Objectives -- (**F**ind, **I**dentify, **S**elect and **O**btain)
- Metadata Application Profile (MAP)
- DOI requirements
- Key elements in the OOI environment
- Related information & relationship (cruise information, calibration information, etc.)
- Use relatively fixed context (array, platform, instrument) to generate metadata for more dynamic content (data streams)

Pioneer (Array)



PDF

- Diagram (112.91 kb)
- Map (463.82 kb)

Citation & Export

Related Resources (13)

Staff View

Description

Title Pioneer (Array)

Research genre Coastal; Field

Type of item Collection

Date(s) of creation 2013-11-01

Date of deployment

Abstract or summary The first phase of the Pioneer Array deployment took place in November 2013, with the installation of three moorings – a surface mooring at the central site, a profiler mooring at the upstream inshore site and a profiler mooring at the upstream offshore site. In April 2014, during the second deployment phase, five Wire-Following Profile Moorings were deployed at the Pioneer Array with three gliders. The third and final deployment phase is scheduled for Fall 2014.

The Pioneer Array is located off the coast of New England, south of Martha's Vineyard. The Continental Shelf-Slope area off the New England coast is a highly productive area and one that is located at a dynamic intersection where ocean currents meet in weather-like "fronts," and where nutrients, pollutants, and other properties are exchanged between the coast and the deep ocean. Data from the inshore, shelf area helps to examine exchanges between the shelf and slope and the shelf ecosystem, as well as provide broader insight into the issues of air-sea gas exchange, including Carbon Dioxide.

The backbone of the Pioneer Array will be a frontal-scale moored array with three electro-mechanical (EM) surface moorings and seven profiler moorings. Each surface mooring will incorporate a surface buoy with multiple sources of power generation and multiple surface and subsurface communications systems. Each surface mooring will be anchored by a Multi-Function Node (MFN) on the seafloor. Two MFNs will incorporate docking stations for AUVs. All three MFNs will be capable of supporting multiple onboard (e.g., frame-mounted) sensors as well as external sensor packages connected by wet-mateable connectors. Five profiler moorings with surface expressions for data telemetry will support wire-following profiling packages with a multidisciplinary sensor suite. Two surface-piercing profiler moorings will use a buoyant instrument package capable of profiling from a few meters above the bottom up to the air-sea interface. In order to provide synoptic, multi-scale observations of the outer continental shelf, shelf break region, and continental slope, the moored array will be supplemented by nine mobile platforms – six gliders and three AUVs. The role of the gliders will be to monitor the mesoscale field of the slope sea and outer shelf, resolving rings, eddies and meanders from the Gulf Stream as they impinge on the shelf break front. The AUVs will be the primary tools for resolving cross- and along-front "eddy fluxes" due to frontal instabilities, wind forcing, and mesoscale variability.

Location

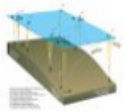
Location of the central mooring: 40.14°N, 70.78°W. Approximate Water Depth: 133 meters at center of moored array

Funder or Sponsor Consortium for Ocean Leadership

Subjects Oceanography--Observations--Remote sensing: Oceanographic research stations--North Atlantic Ocean: 40.14°N, 70.78°W

Rights statement This data has been made openly available by the Ocean Observatories initiative. There are no restrictions on its use. The OOI Program requires an acknowledgement (in publications, conference papers, etc.) from those who use data from OOI, its tools or software. Users are required to specifically acknowledge the National Science Foundation and the COL OOI Program Office when core data/infrastructure is used and individual researchers, groups, or organizations when project specific data is used.

Pioneer (Array)



PDF

- Diagram (112.91 kb)
- Map (463.82 kb)

Citation & Export

Related Resources (13)

Staff View

Related Resources (13)

Hide

1 Ocean Observatories Initiative data (Collection)
2 Pioneer (Central). CP01CNSM, High-Power Surface Mooring (PhysicalObject)
3 Pioneer (Central). CP01CNSP, Surface-Piercing Profiler Mooring (PhysicalObject)
4 Pioneer (Central Inshore). CP02PMCI, Coastal Profiler Mooring (PhysicalObject)
5 Pioneer (Central Offshore). CP02PMCO, Coastal Profiler Mooring (PhysicalObject)
6 Pioneer (Inshore). CP03ISSM, High-Power Surface Mooring (PhysicalObject)
7 Pioneer (Inshore). CP03ISSP, Surface-Piercing Profiler Mooring (PhysicalObject)
8 Pioneer (Mobile). CP05MOAS-AV, AUV (PhysicalObject)
9 Pioneer (Mobile). CP05MOAS-GL, Coastal Glider (PhysicalObject)
10 Pioneer (Offshore). CP04OSPM, Coastal Profiler Mooring (PhysicalObject)
11 Pioneer (Offshore). CP04OSSM, High-Power Surface Mooring (PhysicalObject)
12 Pioneer (Upstream Inshore). CP02PMUI, Coastal Profiler Mooring (PhysicalObject)
13 Pioneer (Upstream Offshore). CP02PMUO, Coastal Profiler Mooring (PhysicalObject)

Description

Title Pioneer (Array)

Research genre Coastal; Field

Type of item Collection

Date(s) of creation 2013-11-01

Abstract or summary The first phase of the Pioneer Array deployment took place in November 2013, with the installation of three moorings – a surface mooring at the central site, a profiler mooring at the upstream inshore site and a profiler mooring at the upstream offshore site. In April 2014, during the second deployment phase, five Wire-Following Profile Moorings were deployed at the Pioneer Array with three gliders. The third and final deployment phase is scheduled for Fall 2014.

The Pioneer Array is located off the coast of New England, south of Martha's Vineyard. The Continental Shelf-Slope area off the New England coast is a highly productive area and one that is located at a dynamic intersection where ocean currents meet in weather-like "fronts," and where nutrients, pollutants, and other properties are exchanged between the coast and the deep ocean. Data from the inshore, shelf area helps to examine exchanges between the shelf and slope and the shelf ecosystem, as well as provide broader insight into the issues of air-sea gas exchange, including Carbon Dioxide.

Pioneer (Central). CP01CNSM, High-Power Surface Mooring

PDF

PDF-1 (139.86 kb)

Citation & Export

Related Resources (12)

Usage Statistics

Staff View

- Data
- Project Documents
- Software
- Instrumentation

Related Resources (12)

Hide

1 3-Wavelength Fluorometer: Wet Labs. BBFL2W ECO TRIPLET-W. (CP01CNSM-00004) (PhysicalObject)
2 Absorption Spectrophotometer: WET Labs. AC-S. (CP01CNSM-00004) (PhysicalObject)
3 Bulk Meteorology Instrument Package: Star Engineering. ASIMET. (CP01CNSM-00004) (PhysicalObject)
4 Bulk Meteorology Instrument Package: Star Engineering. ASIMET. (CP01CNSM-00004) (PhysicalObject)
5 CTD Pumped: Sea-Bird. SBE 16plusV2. (CP01CNSM-00004) (PhysicalObject)
6 Dissolved Oxygen Stable Response: Aanderaa. 4831. (CP01CNSM-00004) (PhysicalObject)
7 Nitrate: Satlantic. ISUS. (CP01CNSM-00004) (PhysicalObject)
8 pCO2 Air-Sea: Pro-Oceanus. pCO2-pro. (CP01CNSM-00004) (PhysicalObject)
9 Pioneer (Array) (Collection)
10 Single Point Velocity Meter: Nortek. Aquadopp 300M. (CP01CNSM-00004) (PhysicalObject)
11 Spectral Irradiance: Satlantic. ICSW. (CP01CNSM-00004) (PhysicalObject)
12 Surface Wave Spectra: Axys Technologies. TRIAXYS. (CP01CNSM-00004) (PhysicalObject)

Description

Title Pioneer (Central). CP01CNSM, High-Power Surface Mooring

Research genre Instruments; Field

Type of item PhysicalObject

Abstract or summary Platform assembly containing data collection instruments/sensors on the Pioneer Array, a network of platforms and sensors operating on the continental shelf and slope south of New England in the Atlantic Ocean.

The Pioneer Central Surface Mooring is located on the Continental Shelf, approximately 130 meters deep. The Continental Shelf-Slope area off the New England coast is a highly productive area and one that is located at a dynamic intersection where ocean currents meet in weather-like "fronts," and where nutrients, pollutants, and other properties are exchanged between the coast and the deep ocean. Data from the shelf area help to examine exchanges between the shelf and slope and the shelf ecosystem, as well as provide broader insight into the issues of air-sea gas exchange, including Carbon Dioxide.
Like other coastal moorings, the Pioneer Central Surface Mooring is specifically designed to examine coastal-scale phenomena and withstand the challenging conditions of shallow coastal environments, including large tidal fluctuations. The Surface Mooring contains a surface buoy floating on the sea surface and instruments located at fixed depths through the water column, and an anchor on the seafloor. The surface buoy provides a platform on which to secure surface instruments, allowing for the collection of data in the air and in the water, as well as an antenna to transmit data to shore via satellite.

Funder or Sponsor Ocean Observatories Initiative

Sea-Bird SBE 52MP

PDF

PDF-1 (418.66 kb)

Citation & Export

Related Resources (3)

Usage Statistics

Staff View

Description

Title Sea-Bird SBE 52MP

Name Sea-Bird

Date Issued

Other Date 2014-04-07- 2014-04-19 (date deployed)

Subject Ocean temperature--North Atlantic Ocean, Electric conductivity, Salinity--North Atlantic Ocean, Detectors--North Atlantic Ocean, Seawater--Density--Observations, Oceanographic observations

Description The SBE 52-MP is a Conductivity, Temperature, Depth (pressure) sensor, designed for moored profiling applications in which measurements are made from from a device that travels vertically beneath a buoy, or from a buoyant package that is winched up and down from a bottom-mounted platform.

Genre Instruments, CTDPF

Persistent URL http://dx.doi.org/doi:10.7282/T33R0VM1

Data Life Cycle Event(s)

Type: Calibration (Initial)
Label: Sensor calibrated and tested.
Date: approximately 2014-04
Detail: CC_latitude: 40.22686667 CC_longitude:-70.88893333
Service provider: Lemmond, Peter (Woods Hole Oceanographic Institution)
Service provider: Simms, Dave (Seabeam)
Name: Spreadsheet of calibration settings (Omaha_Cal_Info_CP02PMCI_00001_v5x)
Reference: http://dx.doi.org/doi:10.7282/T3V69MCQ

Collection Pioneer Array

Pioneer Array. Sea-Bird SBE 52MP. 2015-07-23/2015-08-07. Raw data.

HYPERLINK

Raw data down...

Citation & Export

Related Resources (1)

Statistics

Staff View

Description

Title Pioneer Array. Sea-Bird SBE 52MP. 2015-07-23/2015-08-07. Raw data.

Type of item Dataset

Date(s) of creation 2015-09-30T00:00:00-04:00

Abstract or summary This data was created by the Ocean Observatories Initiative (OOI) project. The Ocean Observatories Initiative (OOI) project has constructed an interactive, globally distributed, and integrated network of ocean nodes that create an observatory enabling transformational, complex, interdisciplinary ocean science. More information about the OOI can be found at http://oceanobservatories.org/

Funder or Sponsor Ocean Observatories Initiative

Research domain Science: Marine Science

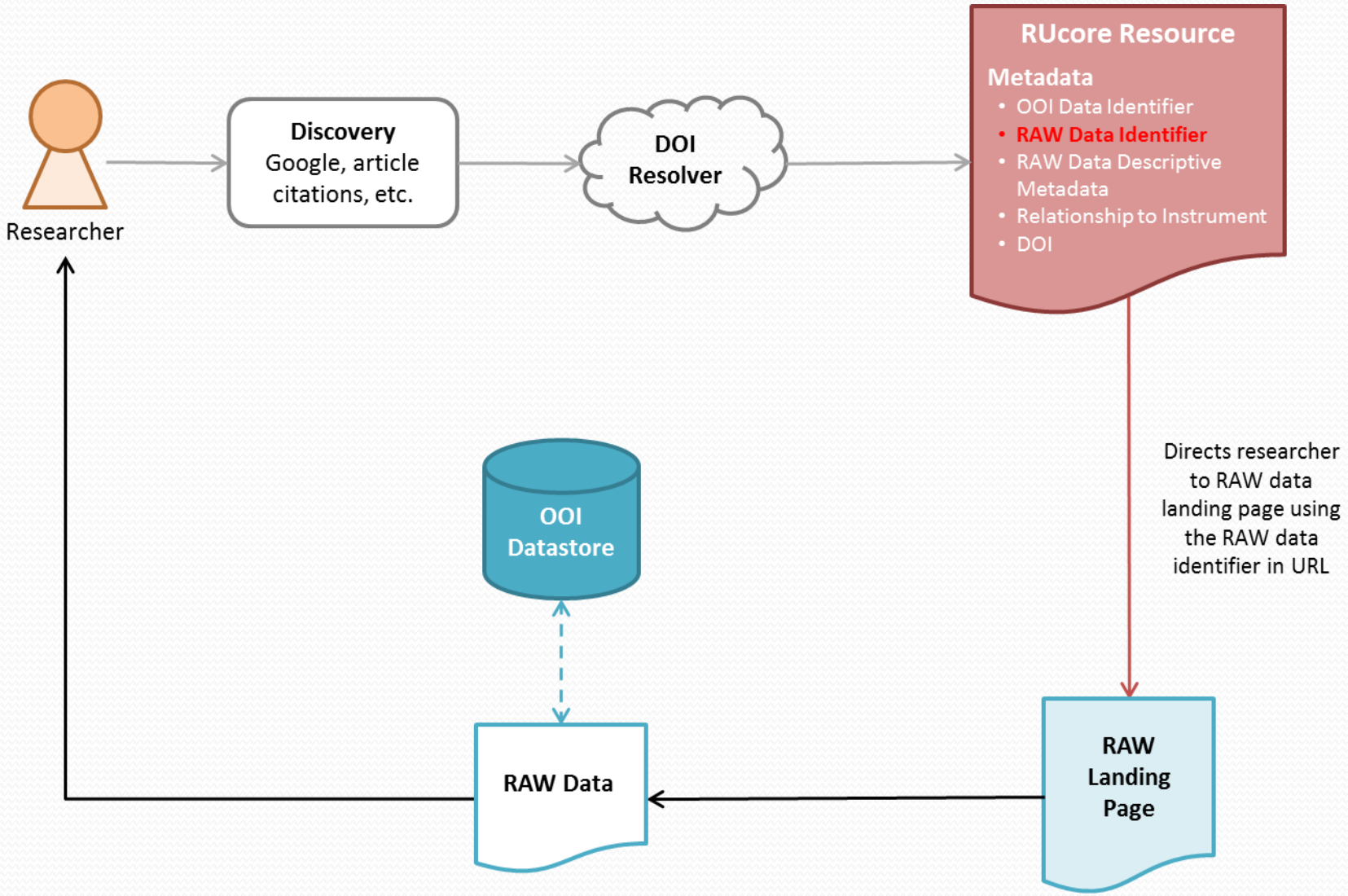
Subjects Oceanography--Observations--Remote sensing: Oceanographic research stations--North Atlantic Ocean: Electric conductivity: Seawater--Density--Observations: Salinity--North Atlantic Ocean

Rights statement This data has been made openly available by the Ocean Observatories initiative. There are no restrictions on its use.

Citation URL http://dx.doi.org/doi:10.5072/FK22R3TD8D

Data Citation Resolution Workflow – Raw Data

9/22/2015.v2 – Mills

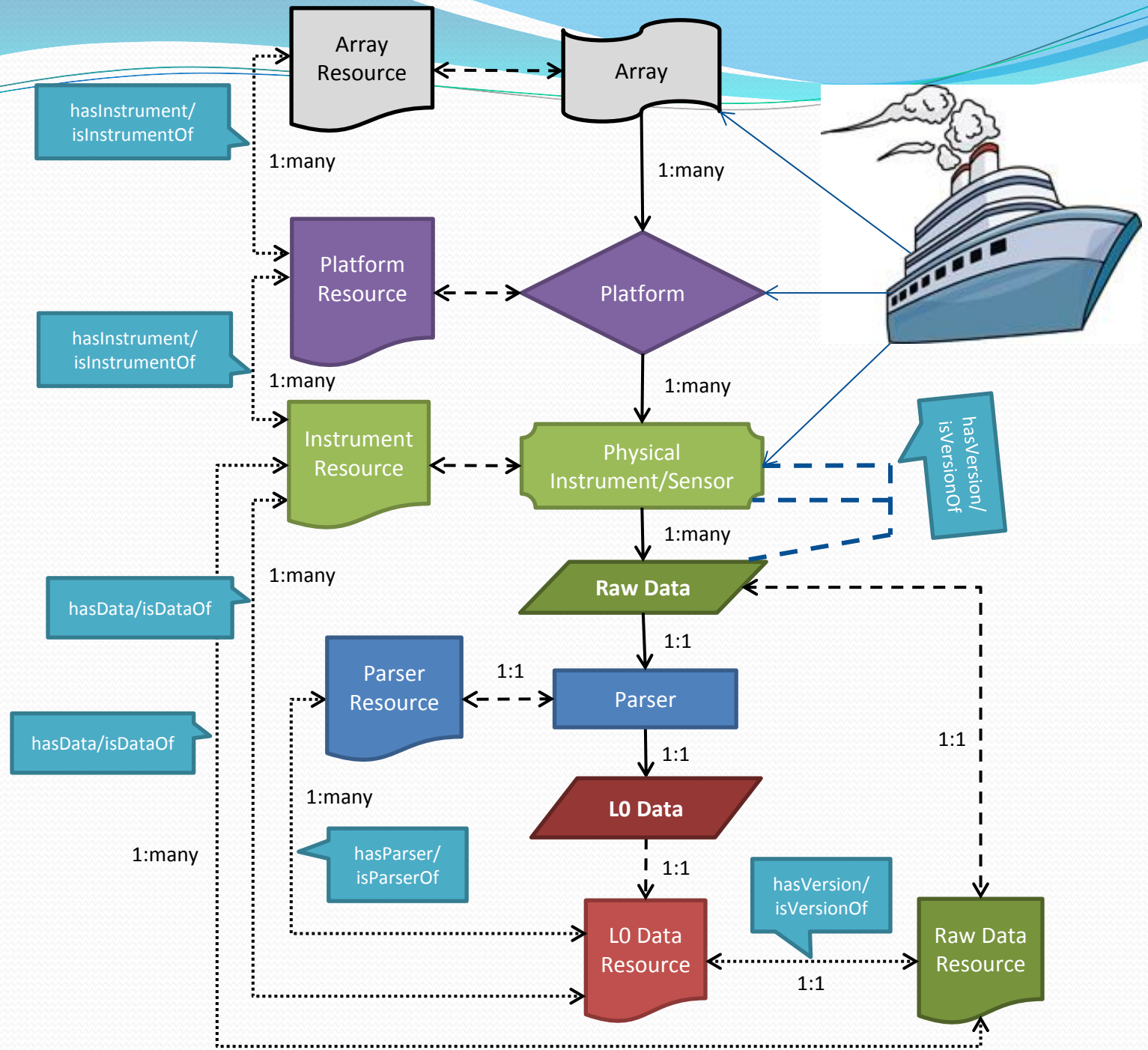


Versioning of Data

When major context changes, a data stream is ended and a new version created:

- Instrument is recalibrated, repaired or replaced
- Instrument is relocated on the same platform (where geographic difference does not mean new data)
- Instrument is out of commission a significant length of time.

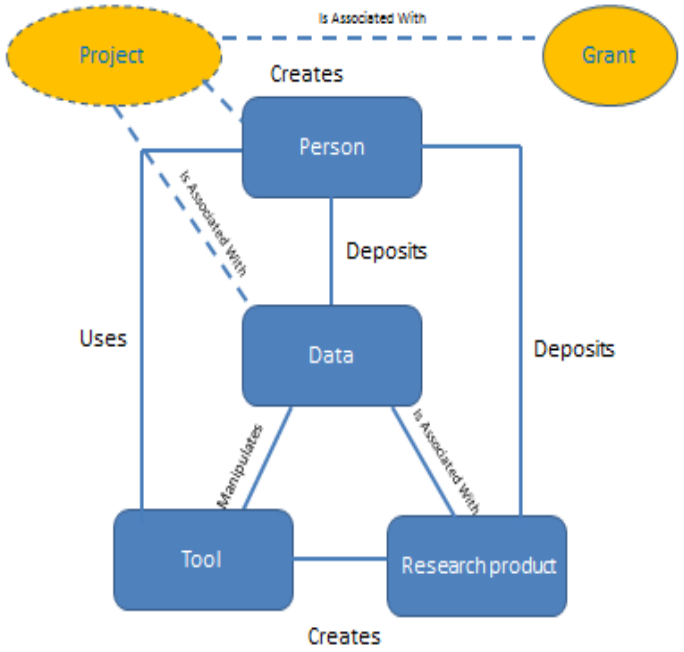
- Needed a strategy for automatic notification of condition changes to automatically generate new versions of platforms, equipment, data


Array Resource
Array
hasInstrument/ isInstrumentOf
1:many
1:many
Platform Resource
Platform
hasInstrument/ isInstrumentOf
1:many
1:many
Instrument Resource
Physical Instrument/Sensor
hasVersion/ isVersionOf
1:many
1:many
Raw Data
hasData/isDataOf
1:1
Parser Resource
1:1
Parser
hasData/isDataOf
1:1
1:many
L0 Data
1:many
hasParser/ isParserOf
1:1
hasVersion/ isVersionOf
L0 Data Resource
Raw Data Resource
1:1

Virtual Data Collaboratory

- NSF CIF21 DIBBS EI #1640834 Virtual Data Collaboratory: A Regional Cyberinfrastructure for Collaborative, Data Intensive Science (of Rutgers, Penn State, et al)
  - Scientific Workflow Support
  - Tool rich. Use tools, create intermediate products, store, reuse, make available to others
- An integration framework to enable VDC to interoperate with other large data repositories
  - APIs enable VDC users to discover resources outside the VDC, using the powerful search and browse capabilities of the external repository, but to leverage the resources within the VDC

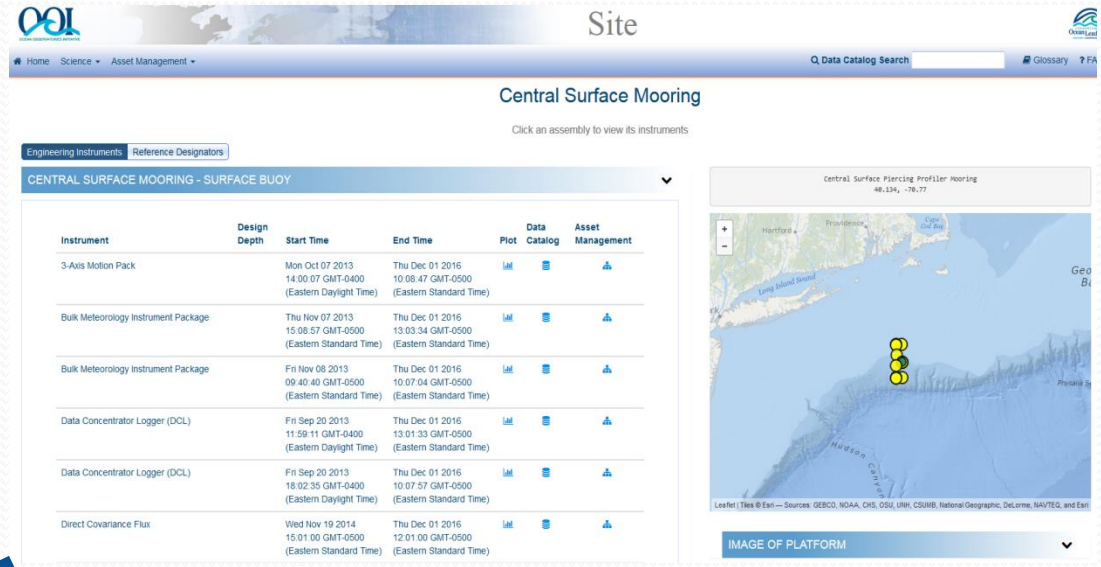
Linked Data Discovery Layer



DOI retrieves data from VDC

Next Requestor of data 2016-12-01 thru 2016-12-31

Virtual Data Collaborative



DOI returned to OOI

Data pulled into VDC. Metadata auto-generated and mapped. DOI generated

VDC participant requests data 2016-12-01 through 2016-12-31