Overview

1. Team Structure & Responsibilities
2. Data Flow & Products
3. Data Review
4. Communications
5. Improvements
6. Conclusions
Data Flow Example: Pioneer Profiler

Ocean

CTDPF DOFST FLORT PARAD VEL3D

Profiler Controller → Inductive Modem

Platform Controller → Telemetry → Iridium

OMC Platform Shore Server

rsync

OMC Data Server

WHOI

Rutgers

Data Files from CGSN

RU Acquisition Point Server → Dataset Agent Driver → uFrame Database → Data Product Algorithm → GUI → User

CTDPF DOFST FLORT PARAD VEL3D

RU Acquisition Point Server

Dataset Agent Driver

uFrame Database

Data Product Algorithm

GUI

User
Current Data Processing Flow

**Instrument Preparation**
- Add new Asset info and UID to bulk load
- Add new Calibration Info

**Pre-Cruise**
- Submit Pull Request
- Add new Deployment Info
- Add new Cruise Info

**Post-Cruise Telemetered & Streamed Data**
- Revise Deployment Info
- Revise Cruise Info
- Update Pull Request
- Submit Redmine Ticket

**Post-Cruise Recovered Data**
- Upload Recovered Data
- Submit Redmine Ticket

**CI Data Team**
- Review and Merge PR
- Review Deployment & Cruise Info
- Prepare Ingestion Info

**CI System Team**
- Synchronize on a regular basis

**Notes**
- Sheet
- GitHub Commit
- Action
- Redmine Ticket
OOI Data Sheets (Calibration, Ingestion, Deployment)

Calibration Sheet

Ingestion Sheet

Deployment Sheet

1 per instrument

0-2 per platform

1 per deployment
Data Types

• Telemetered Data
  o Data received through a transmission medium over distance (e.g. surface buoy to satellite, glider to satellite, acoustic modem); may be decimated

• Recovered Data
  o Data downloaded directly from a recovered instrument or data logger after the instrument has been recovered.

• Streamed Data
  o Data received via transmission over electro-optical cable. Streaming data are provided at full temporal resolution and near-real time.

• Shipboard Data
  o Shipboard data and water samples collected during OOI expeditions.

• Metadata
  o Info about the data record (e.g., time & location of collection, unique source & record description identifier, instrument serial #, etc.). OOI metadata follows the CF1.6 standard, with additional types and fields specific to OOI as necessary.
OOI Data Product Levels

- **Raw data**: The datasets as they are received from the instrument
  - May contain multiple L0, L1, or L2 parameters, data for multiple sensors, and be in native sensor units
  - Always persisted and archived by the OOI
  - **Example**: format 0 binary file from an SBE-37IM on a Global Flanking Mooring.

- **Level 0 (L0)**: Unprocessed, parsed data parameter that is in instrument/sensor units and resolution
  - Sensor by sensor (unpacked and/or de-interleaved) and available in OOI supported formats (e.g., NetCDF)
  - Always persisted and archived by the OOI
  - **Example**: SBE-37IM Temperature portion of the hex string

- **Level 1 (L1)**: Data parameter that has been calibrated and is in scientific units
  - QC may be applied at this level, utilizing simple automated techniques or human inspection
  - Actions to transform Level 0 to Level 1 data are captured and presented in the metadata of the Level 1 data
  - **Example**: SBE-37IM Temperature converted from hex to binary and scaled to produce degrees C

- **Level 2 (L2)**: Derived data parameter created via an algorithm that draws on multiple L1 data products
  - Products may come from the same or from separate instruments
  - Data from all relevant instruments will be provided during download
  - **Example**: SBE-37IM Density and Salinity
OOI: Web Portals

Main

OOINet

Education

Cruises

May 2017 OOI FB Meeting