
ESIP's Marine Data Cluster

Introduction and Development of Focus Areas

Jocelyn Elya

Florida State University

jelya@coaps.fsu.edu

Carolina Berys-Gonzalez, Chris Olson, Steve Diggs, Karen Stocks, Joseph Gum
Scripps Institution of Oceanography

2019 ESIP Winter Meeting
Tuesday, January 15, 2019 4:00 pm

**“Marine Data” ~ Data taken over or within
a body of water**

What is the ESIP Marine Data Cluster?

Formed at Summer 2018 ESIP meeting

Thursday Ocean Data Group formed March 2017 as a casual Thursday call

ESIP Marine Data Cluster = Thursday Ocean Data Group as of 11/16/2018



Chair: Jocelyn Elya
FSU/COAPS



Co-chair: Carolina Berys-Gonzalez
SIO/CCHDO



Co-chair: Christopher Olson
SIO/R2R

Mission Statement

“The goal of the Marine Data Cluster is to bring together ESIP members working with data in the marine geosciences to discuss advancements and challenges in their field, and to build relationships to foster future collaborations. This group will explore critical topic areas in informatics such as data management, interoperability, data analytics, data quality, and cloud computing, but specifically related to the marine environment, the members’ shared domain.

The cluster facilitates teleconferences twice a month where members or other cohorts will present their research and solicit feedback from the community. The intent is to connect with participants from other relevant ESIP data clusters and discuss how innovations in their work can be applied to the marine geosciences.”

Mission Statement - “What does it mean???”

“A coalition of the willing to meet online to collaborate, harmonize marine data for the betterment of the community.” -Steve Diggs, mostly (ESIP Vice President)

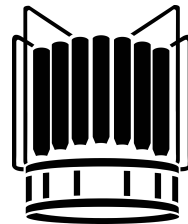
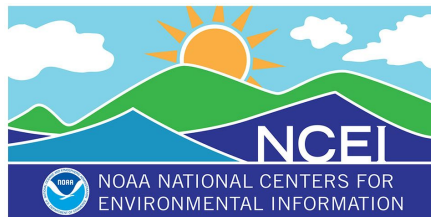
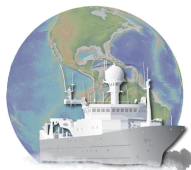
In practice, teleconferences on **2nd and 4th Thursdays of the month**
10:30-11:30 AM Pacific Time / 1:30-2:30 PM Eastern Time

Goals:

- Widespread adoption of netCDF-CF
- Preserving calibration information
- Software versioning and curation
- DOIs as data comes in
- Best practices for managing different data types

A history of Thursday Ocean Data Group

- Starts with CCHDO, Bob Key
- Editing controlled vocabularies for WHP-exchange(?) to deal with non GO-SHIP data coming into CCHDO
- Expands to BCO-DMO, NCEI, R2R, SAMOS, etc as groups have commonalities in problems and solutions
- Invite scientists at large to talk about their work and problems
- Talk about general tech advancements (ERDDAP, ODV, Docker, etc)
- Actionable projects
 - 24 Hz CTD data(Microstructure, LADCP)
 - Submitting vocabulary to CF for approval



HOW STANDARDS PROLIFERATE:

(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)

SITUATION:
THERE ARE
14 COMPETING
STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.

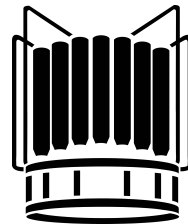
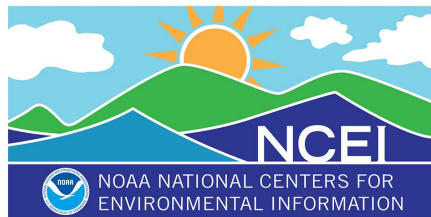
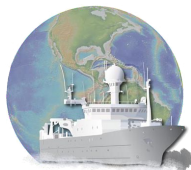


SOON:

SITUATION:
THERE ARE
15 COMPETING
STANDARDS.

A history of Thursday Ocean Data Group

- Starts with CCHDO, Bob Key
- Editing controlled vocabularies for WHP-exchange(?) to deal with non GO-SHIP data coming into CCHDO
- Expands to BCO-DMO, NCEI, R2R, SAMOS, etc as groups have commonalities in problems and solutions
- Invite scientists at large to talk about their work and problems
- Talk about general tech advancements (ERDDAP, ODV, Docker, etc)
- Actionable projects
 - 24 Hz CTD data(Microstructure, LADCP)
 - Submitting vocabulary to CF for approval



Goals for Today's Session

1. Get community feedback on telecon topics
 - a. Introductions from cluster admins: Do the topics we present resonate with you?
 - b. What other topics would you like to see covered?
 - c. How can we collaborate with other ESIP clusters/committees?
2. Determine which 2 or 3 topics to begin addressing in 2019
3. Outline goals for 2019



Jocelyn Elya



jelya@coaps.fsu.edu
[id 0000-0002-2261-1449](https://orcid.org/0000-0002-2261-1449)
[@JocelynElya](https://twitter.com/JocelynElya)

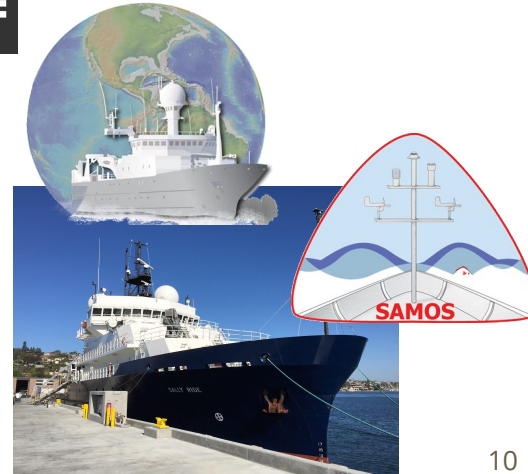
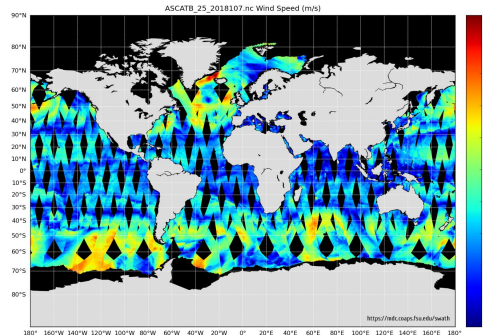
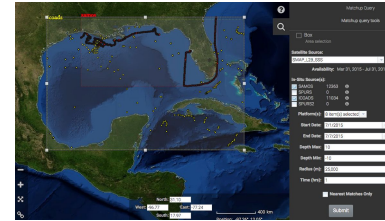
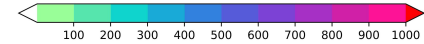
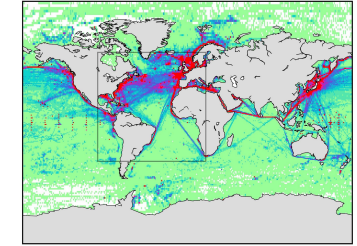
Lead Programmer, Marine Data Center
Center for Ocean-Atmospheric Prediction Studies
Florida State University

Main Projects

- [SAMOS](#) (Shipboard Automated Meteorological and Oceanographic System)
- [OceanWorks](#)
- [R2R](#) (Rolling Deck to Repository)

Potential Cluster Topic/Goal -
Version control for legacy software,
moving to open-source

Number of Observations 1990-01-01 - 1992-12-31





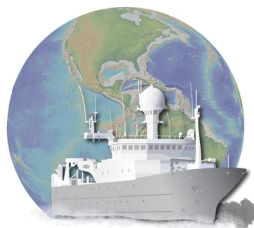
Chris Olson

cjolson@ucsd.edu

Programmer and Data Manager
Geological Data Center
Scripps Institution of Oceanography

Main Projects

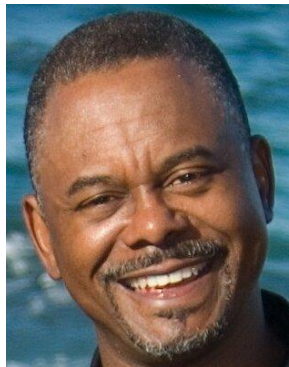
- R2R
- IODP



IODP
INTERNATIONAL OCEAN
DISCOVERY PROGRAM

Potential Cluster Topic -
Creating a standardized way to record
and maintain sensor calibration
information





Steve Diggs



Technical Director, CCHDO
ESIP Vice President



SOOS
SOUTHERN OCEAN
OBSERVING SYSTEM

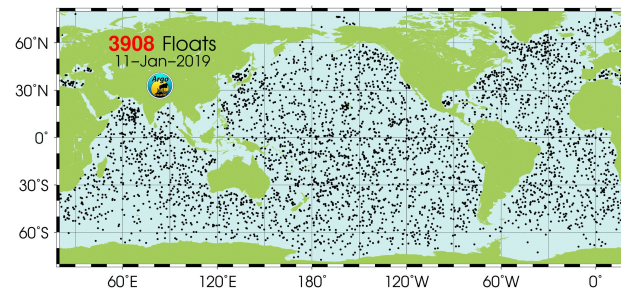
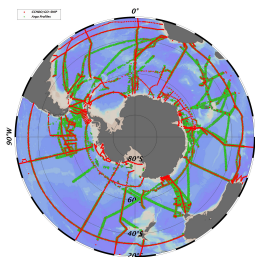


Main Projects

- [CCHDO](#) (CLIVAR and Carbon Hydrographic Data Office)
- [Argo](#) Network
- OceanSITES, Data Management Team
- [SOOS](#) (Southern Ocean Observing System) Chair, Data Management

Potential Cluster Topics

- Full-stack data harmonization across all sub-disciplines of the marine sciences
- Ocean data practices broadly aligned with the rest of geoscience





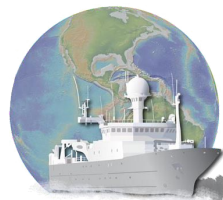
Karen Stocks

Director, Geological Data Center, Scripps Inst. of Oceanography
Director, CCHDO

Main Projects

- [CCHDO](#) (CLIVAR and Carbon Hydrographic Data Office)
- [R2R](#) (Rolling Deck to Repository)
- [IODP](#) (International Ocean Discovery Program; Science Support Office)
- [Discovery Data Studio](#)

Potential Cluster Topic -
Requirements for becoming a
CoreTrustSeal data repository





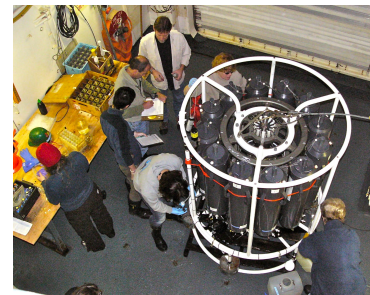
Joseph Gum

jgum@ucsd.edu

ODF Lead Data Analyst
Scripps Institution of Oceanography

Main Projects

- CTD data processing and documentation
- High frequency CTD data archiving
- Shipboard knowledge documentation



Potential Cluster Topic -
Survey, review, and archiving of technical
documentation via IODE -

<https://www.oceanbestpractices.net/>



*Unlocking the mysteries
of the Southern Ocean*

Potential Topics for Cluster Telecons

1. NetCDF-CF compliance
2. Preserving calibration information
3. Legacy software version control, moving to open-source
4. Best practices for managing 24 Hz CTD data
5. DOIs as data come in
6. Parameter reconciliation
7. Continued Docker demos
8. Becoming a CoreTrustSeal data repository
9. Preparing web services for AWS
10. Talks from data managers in other communities (astronomy, biomed, etc)
11. Best practices for software implementation
12. Best practices for models. Managing data, managing code
13. UTF-8 compliance. CSV standards

Topics Poll

<http://etc.ch/bSzb>



Poll Results

<https://directpoll.com/r?XDbzPBd3ixYqg81p0iDMPu6lmL1gCNvREgKXwcwd>

Goals for Select Topics in 2019

1. NetCDF-CF compliance
 - a. Compliance checkers
 - b. Rosetta tool for converting CSV to netCDF-CF files. ERDDAP.
 - c. Parameter reconciliation. Funding for a workshop. CF training
2. Best practices for software implementation
 - a. Version control systems
 - b. Docker demo
 - c. Preparing web services for AWS
 - d. Data servers (ERDDAP, THREDDS)
 - e. Linked data, linking versions
3. Best practices for managing specific marine data types
 - a. 24 Hz CTD data
 - b. Transmissometer and fluorometer data
 - c. NCEI netCDF decision tree
 - d. Storing and sustaining best practice resources

Other Questions

1. How can we engage with other clusters/committees within ESIP and apply the work they are doing to our field?
2. What will we produce?
 - a. Collection of existing standards
 - b. Best practices for different data types
3. Should we pursue funding?
 - a. EarthCube NSF solicitation, Research Coordination Network (RCN):
<https://www.nsf.gov/pubs/2016/nsf16514/nsf16514.htm>

Join Us!

Email List Sign Up: <http://bit.ly/esipmarineemail>

**Teleconferences 2nd and 4th Thursdays of the month
10:30-11:30 AM Pacific Time / 1:30-2:30 PM Eastern Time**

Cluster Wiki: <http://wiki.esipfed.org/index.php/MarineData>

Other Links

- [Thursday Ocean Data Group old website](#)
- [R2R](#)
- [CCHDO](#)
- [FSU COAPS Marine Data Center](#)