CIOOS ATLANTIC

REGIONAL ASSOCIATION OF THE
CANADIAN INTEGRATED OCEAN OBSERVING SYSTEM

COINAtlantic

COASTAL AND OCEAN INFORMATION NETWORK

CIOOS Atlantic

Data Discovery and Contribution

Jeff Cullis, Technical Director, jcullis@cioosatlantic.ca
Alexi Baccardax Westcott, Engagement Specialist, alexi.westcott@cioosatlantic.ca
Shen Molloy, Engagement Specialist, shen.molloy@cioosatlantic.ca

March 9, 2022

Funded by : MEOPAR | Fisheries and Oceans Canada | Pêches et Océans Canada | Tula Foundation

Welcome

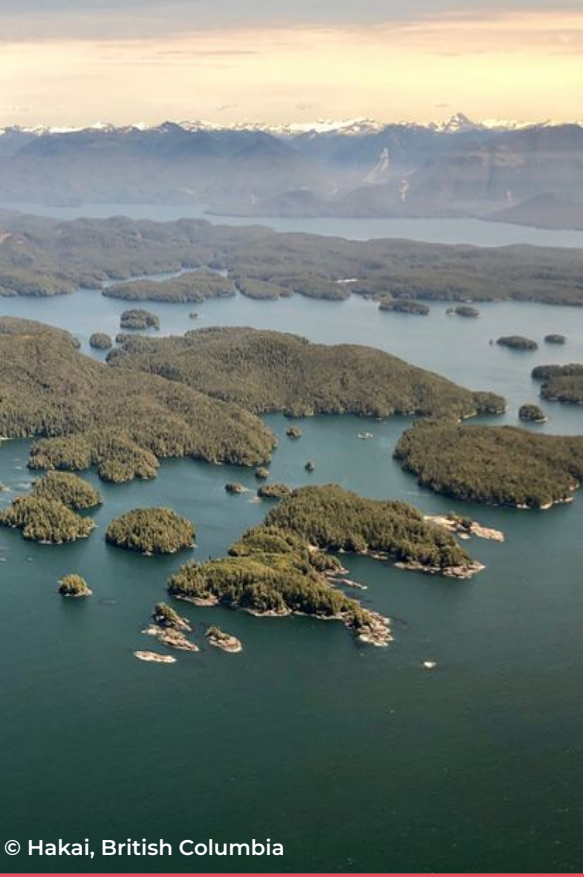

© Hakai, British Columbia

Table of contents

1. About CIOOS

2. Essential Ocean Variables

3. Getting data out

4. Getting data in

5. Questions

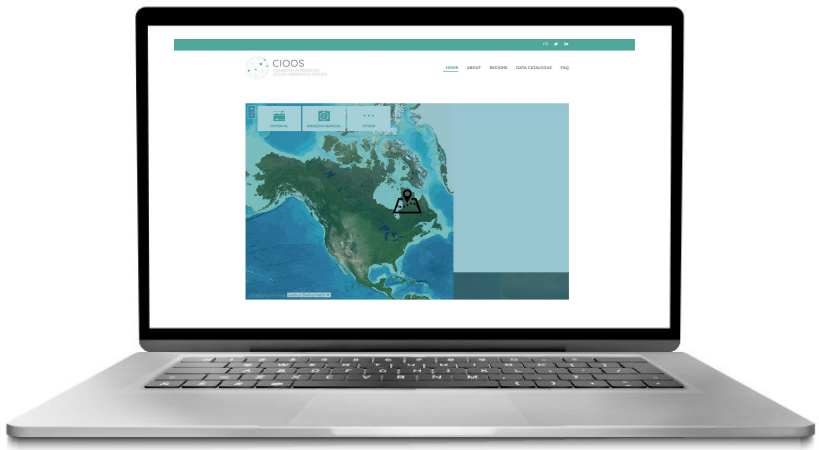
About CIOOS

Ocean Data For Our Ocean Future

CIOOS PACIFIC
REGIONAL ASSOCIATION OF THE
CANADIAN INTEGRATED OCEAN OBSERVING SYSTEM

SLGO.ca
St. Lawrence
Global Observatory

CIOOS ATLANTIC
REGIONAL ASSOCIATION OF THE
CANADIAN INTEGRATED OCEAN OBSERVING SYSTEM



CIOOS works across the country and across sectors to unite the knowledge, expertise and digital infrastructure of Canada's ocean observing community.

CIOOS ATLANTIC

REGIONAL ASSOCIATION OF THE
CANADIAN INTEGRATED OCEAN OBSERVING SYSTEM

OCEAN FRONTIER INSTITUTE

OCEAN TRACKING NETWORK

COINAtlantic
COASTAL AND OCEAN INFORMATION NETWORK

MARINE INSTITUTE

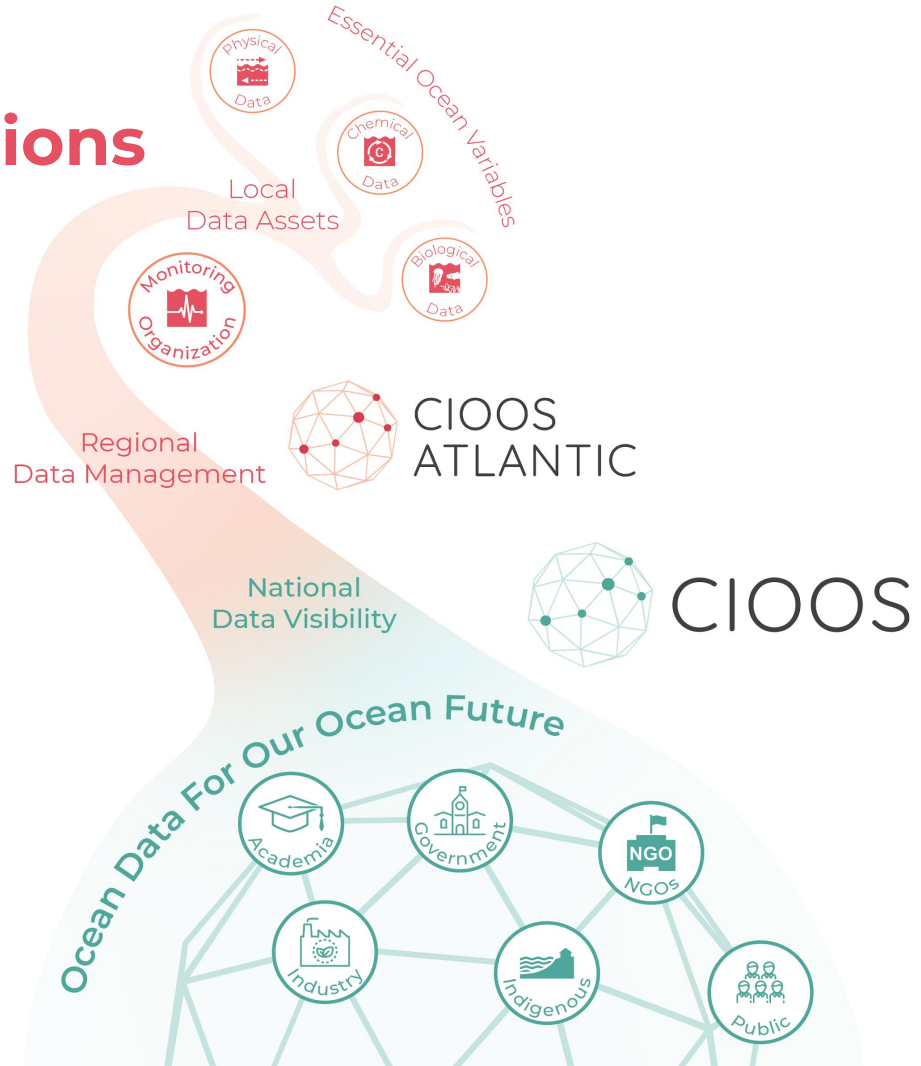
Memorial
University of Newfoundland

DALHOUSIE UNIVERSITY

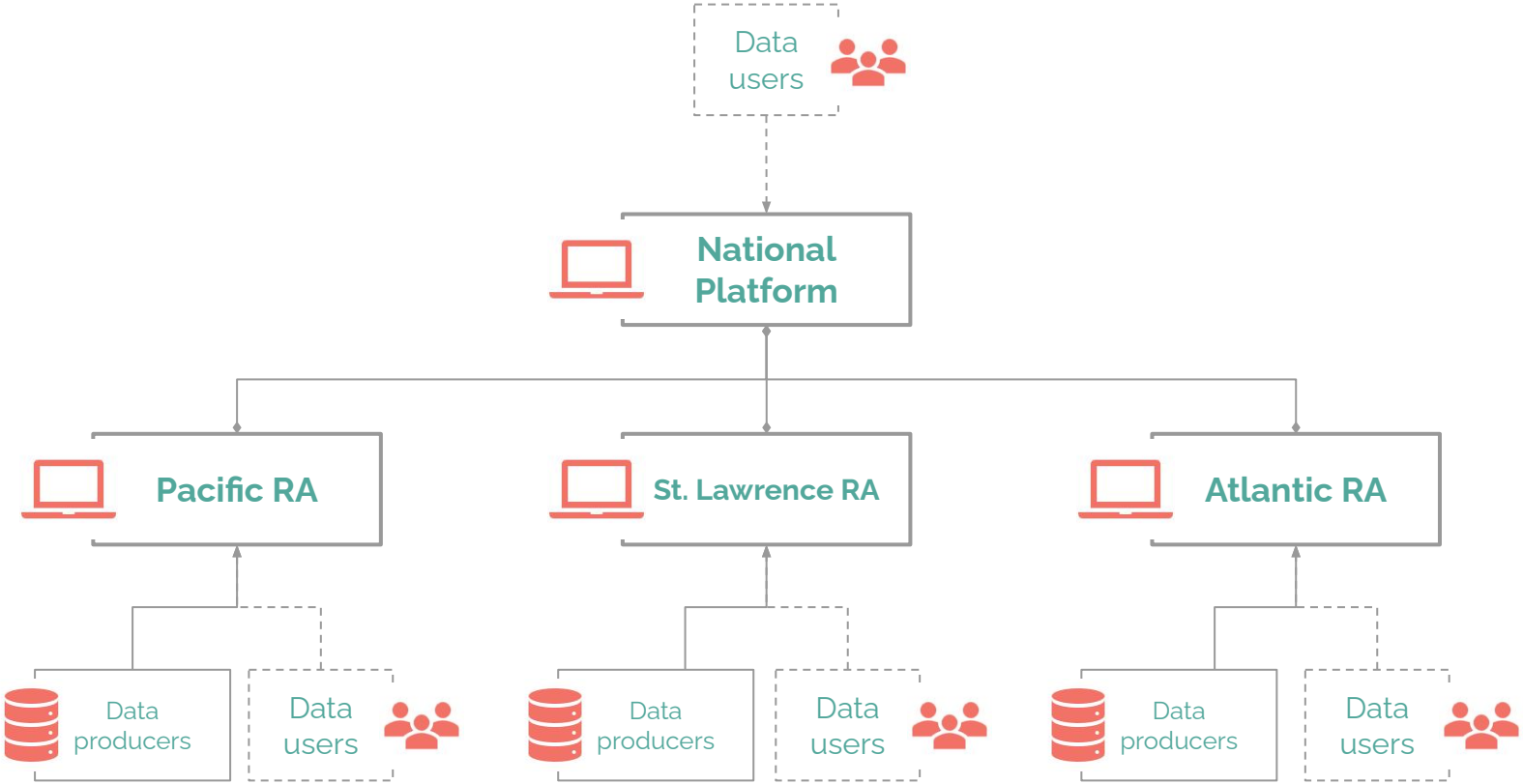
Role of Regional Associations

- Engage and work with regional data contributors (all sectors)
- Meet regional data users
- Foster participation in CIOOS
- Continuous communication in CIOOS

More data =
more robust information system



CIOOS' Structure



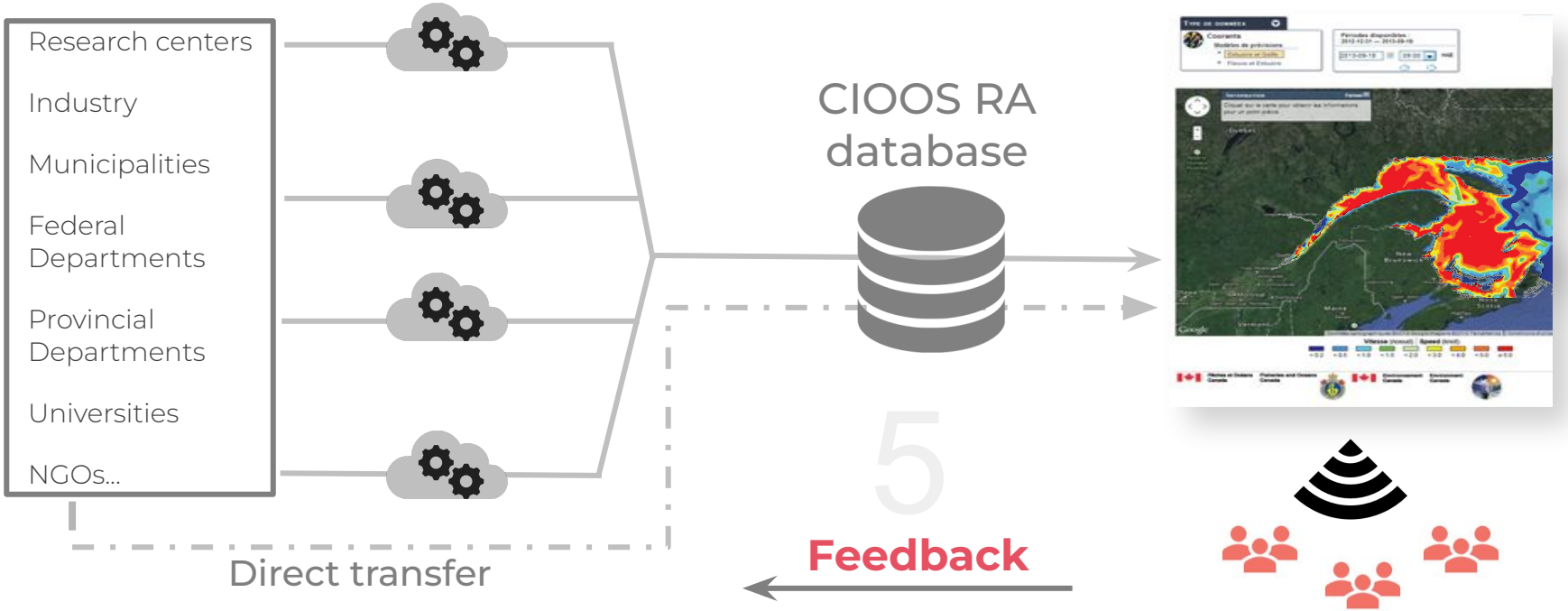
How CIOOS works

1 » Databases of partners

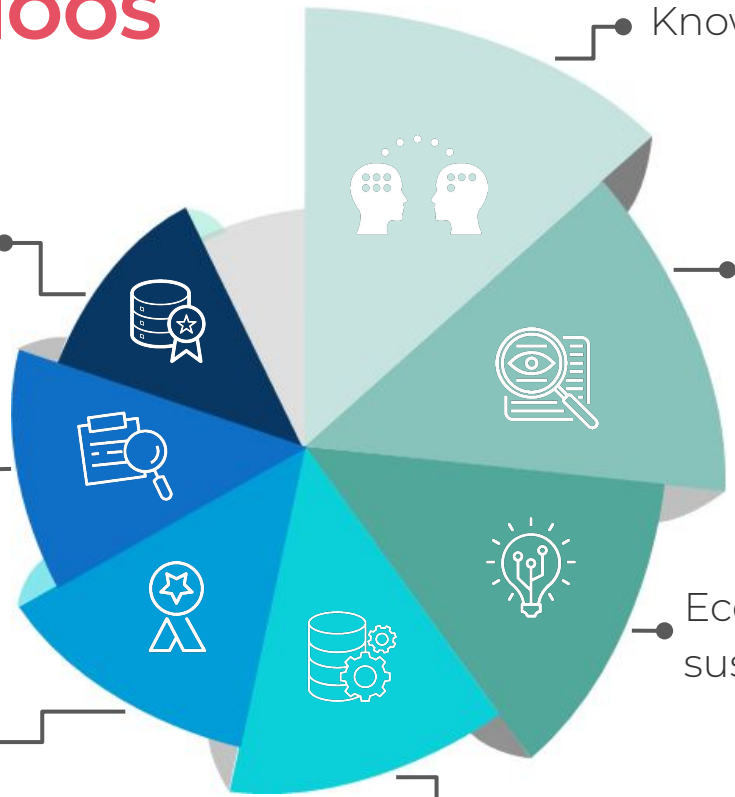
2 » Web Services

3 » Integration and hosting

4 » Open access and visualisation tools



Benefits of CIOOS



Knowledge exchange

Increased transparency

Economic innovation, sustainable development

Management of ocean observing data

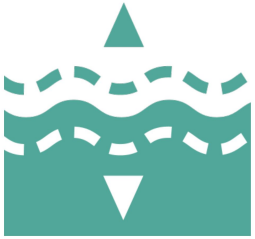
Regional, national and international recognition

Discovery, access, and visualization

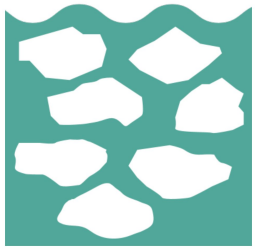
High-quality coastal and ocean data

Essential Ocean Variables (EOVs)

CIOOS EOVs - Physical



Sea Level Rise

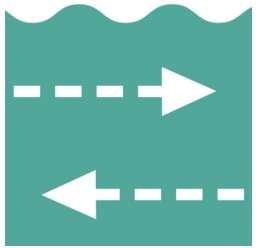


Sea Ice

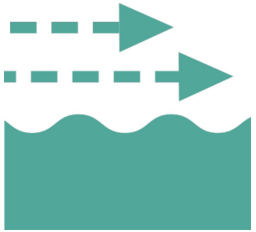
Sea State

Temperature

Salinity



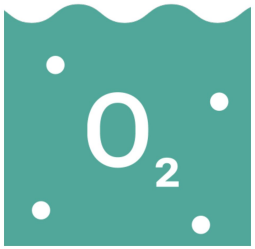
Currents



Ocean Surface Stress

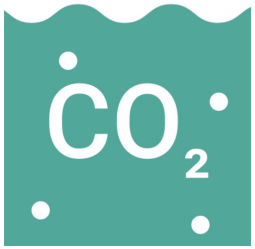
Ocean Surface Heat Flux

CIOOS EOVs - Biogeochemical



Oxygen

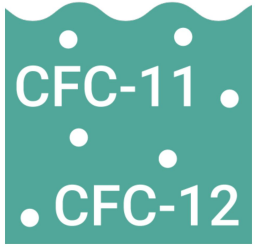
Nutrients



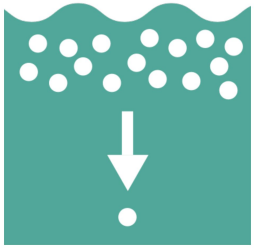
Inorganic Carbon

Dissolved Organic Carbon

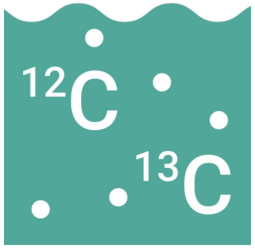
Nitrous Oxide



Transient Tracers

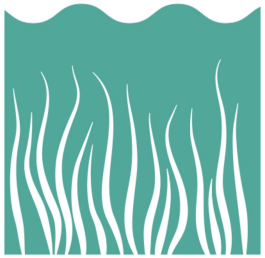


Particulate Matter



Stable Carbon Isotopes

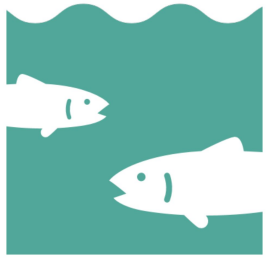
CIOOS EOVs - Biology and Ecosystems



Seagrass

Microalgae

Hard Corals



Fish

Birds and Mammals

Phytoplankton

Zooplankton

Invertebrate (*emerging)

Microbe (*emerging)

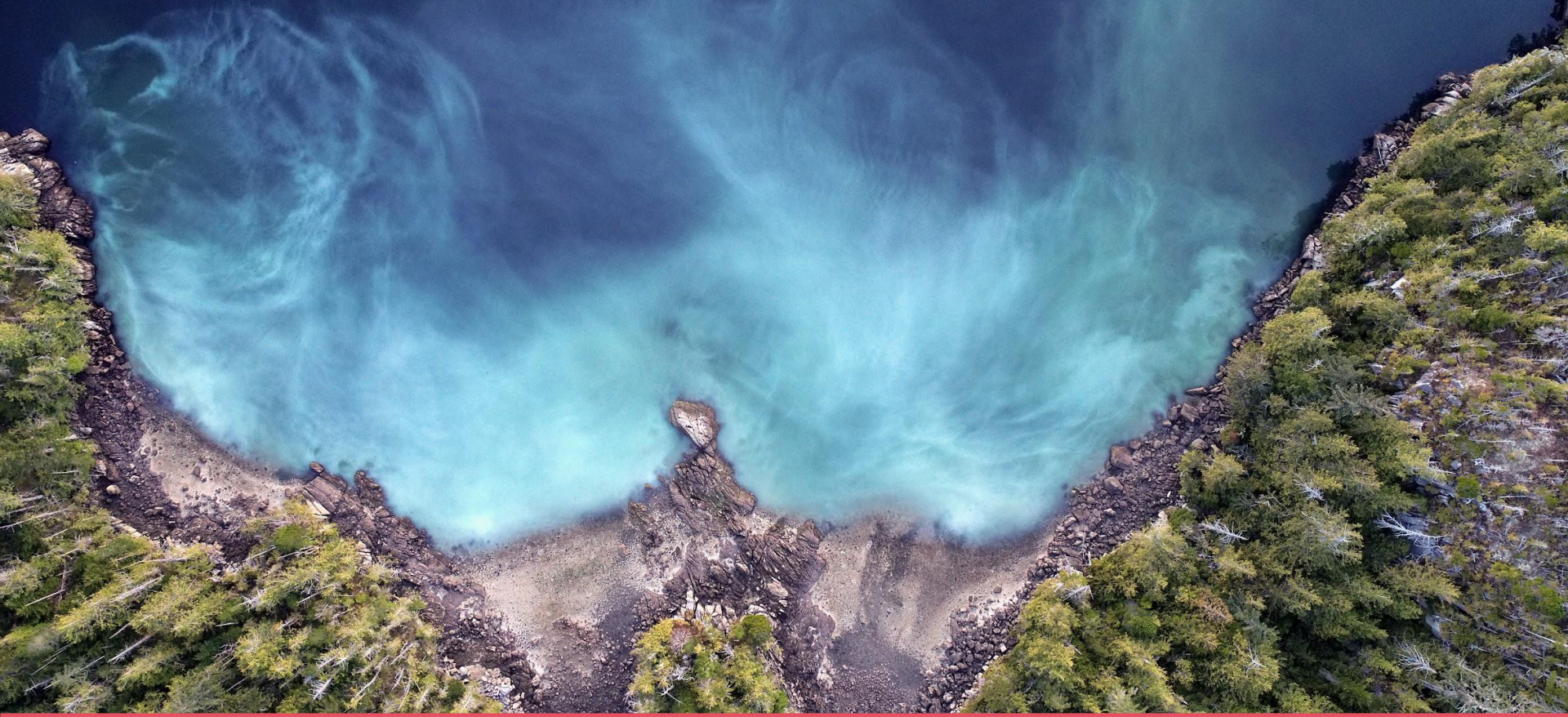
CIOOS EOVs - Cross disciplinary

Marine Debris (*emerging)

Ocean Colour

Ocean Sound

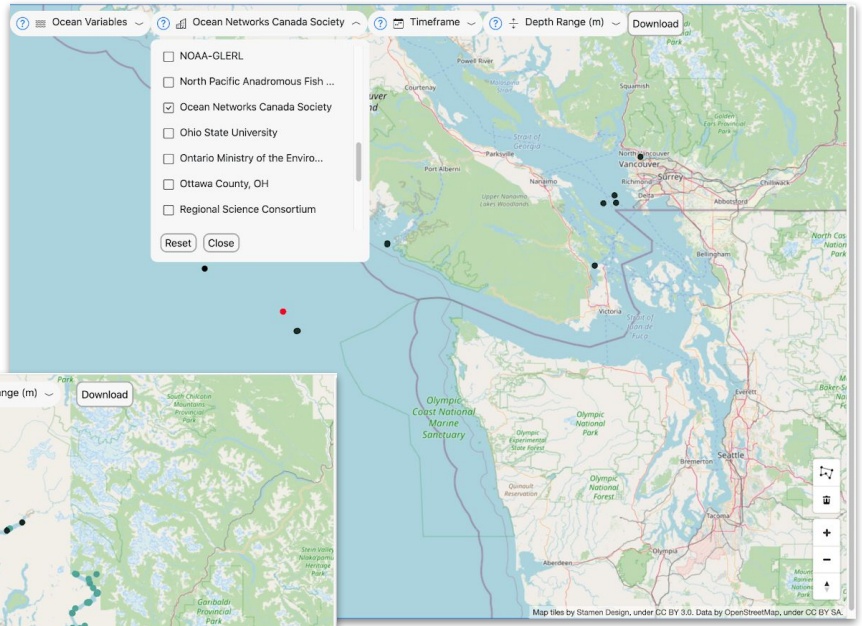
Questions?

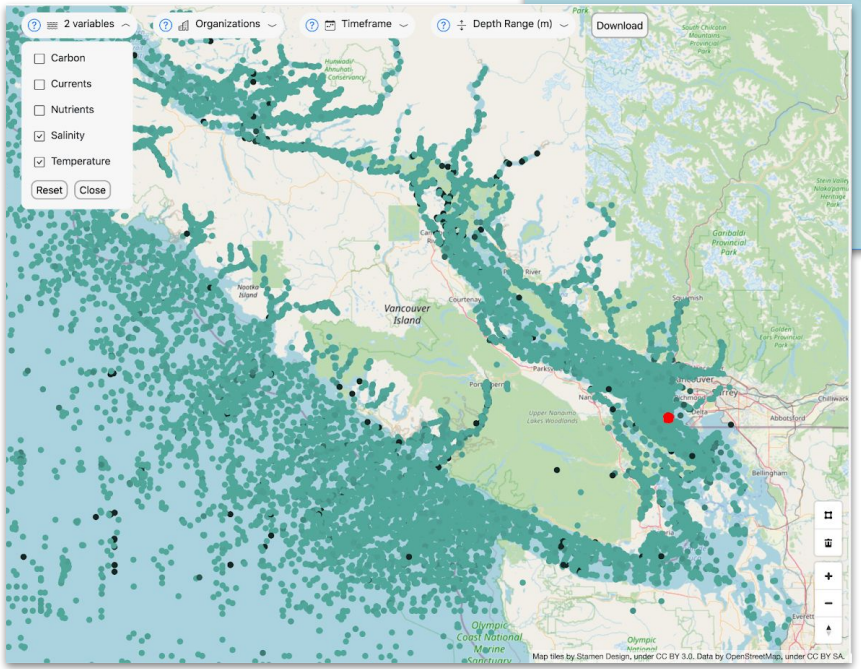


Getting Data Out

Data Explorer - Coming Soon!

A custom data discovery tool that searches and filters across datasets for personalized delivery





New Data Explorer Interface
(work in progress)

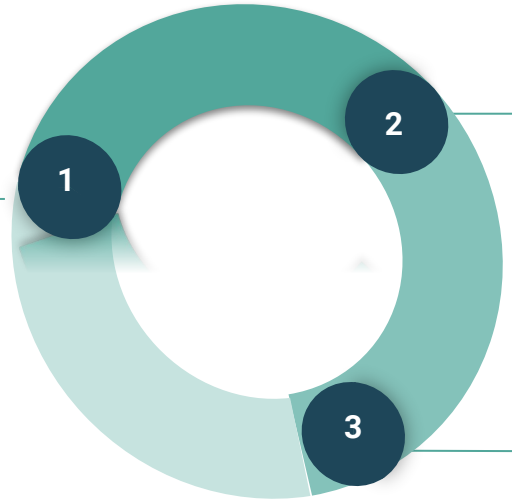
Getting Data In

Getting Data In

1. Engagement Specialist

Your initial point of contact with CIOOS who will help guide you through the process of integrating your data & metadata. They will also liaise with the technical team

info@cioosatlantic.ca



2. Data Specialist

The person who will work in tandem with you and your engagement specialist to bring your data & metadata into CIOOS

3. Review and Iterate

Once your data & metadata have been integrated it will be open for review and refinement before being made public

Initial Meeting

Dataset

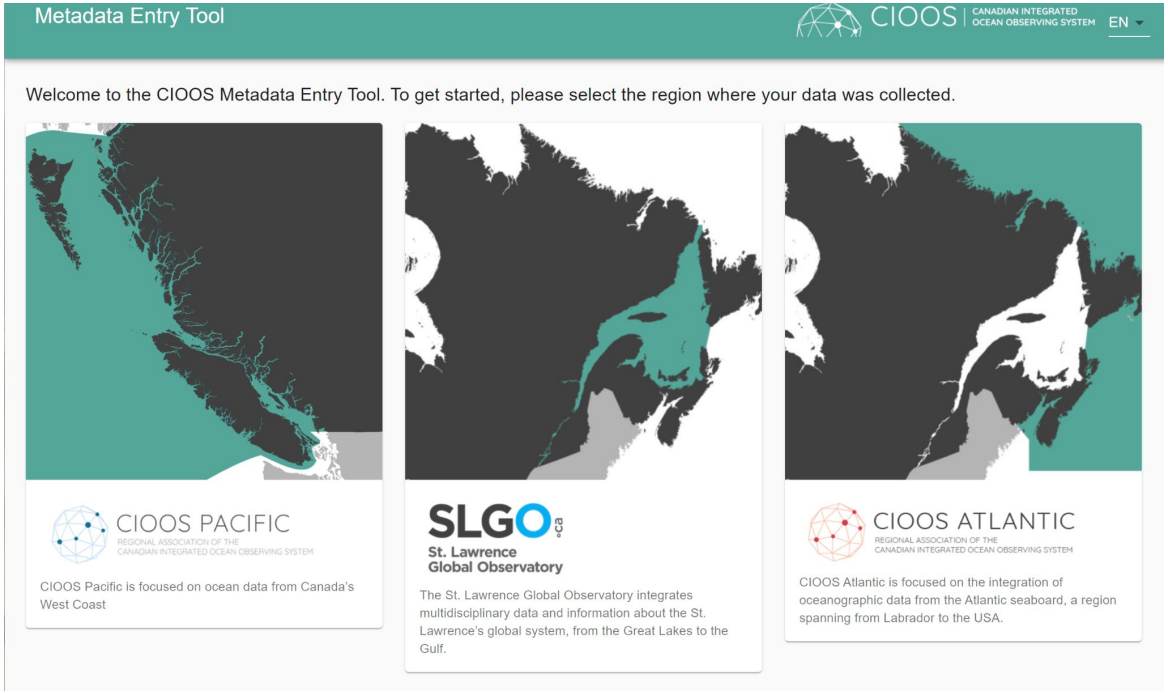
- Size, location, format
- How can CIOOS members access the data
- Completed vs. real-time
- Data licence

Catalogue entry

- Title, summary, EOVs
- Who to contact/attribute
- Dataset identification
- Geospatial
- Platform/instrument

Metadata Entry Tool

- A web based tool for adding datasets to regional metadata catalogues
  - Ensures required fields are completed and does automatic translations
- Added features for region selection, bilingual translations

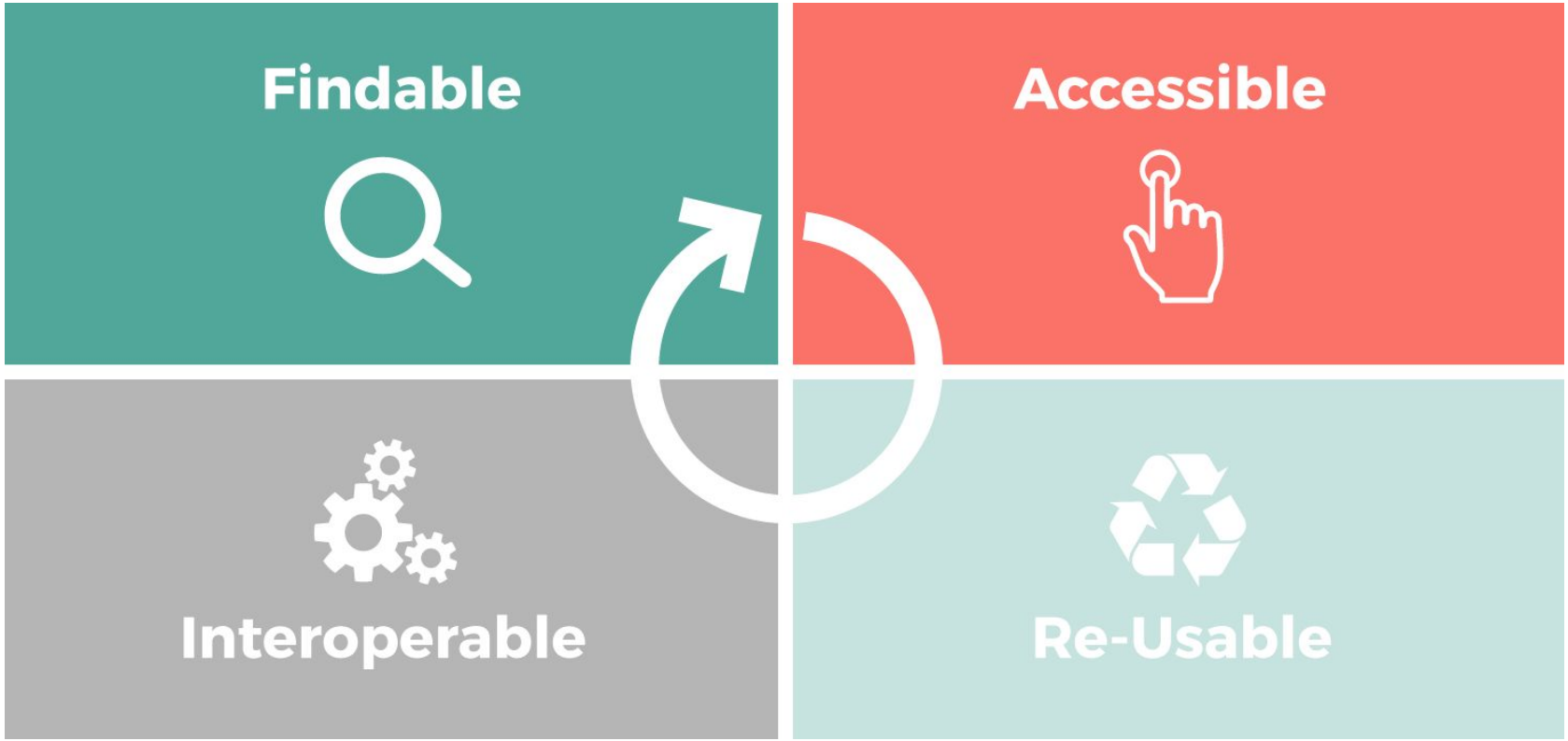


Data Management

Data Management and Integration Overview

- Making your (physical and biogeochemical) data FAIR
- Raw data access
- Data interoperability
- Input data formats
- Standardizing data for interoperability
- Climate and Forecast Conventions and compliance checks
- Data transform and integration examples
- The CIOOS Atlantic data integration workflow
- Next steps once your data is available in the catalogue and ERDDAP
    - Persistent identifiers, dData archival, Data QC
    - API data access and visualization

FAIR Principles



Raw Data Access

- Need a public URL of the dataset in its original form from the data provider
- Ideally the data provider can provide a public URL or API access
- Depending on dataset size, CIOOS Atlantic infrastructure may be able to host it
    - It would be stored as raw files in a simple public web accessible folder (WAF)

Towards FAIR Data: Interoperability

- Data Interoperability
  - Ensure that the data you create is structured such that it can be sensibly combined and aggregated together with other datasets.
- The importance of interoperability
  - Example: different scientists all working with sea water temperature might record it as:
    - ‘Ocean temp’, ‘Temperature’, ‘Water temperature’, ‘H20 temp’, ...
  - The units could be recorded using e.g. ‘C’, ‘degrees C’, Celsius, Fahrenheit ...
  - Combining data from multiple sources now requires:
    - Searching and finding data variables with many different possibilities for temperature
    - Checking whether the units make sense and can all be converted, and then converting them.

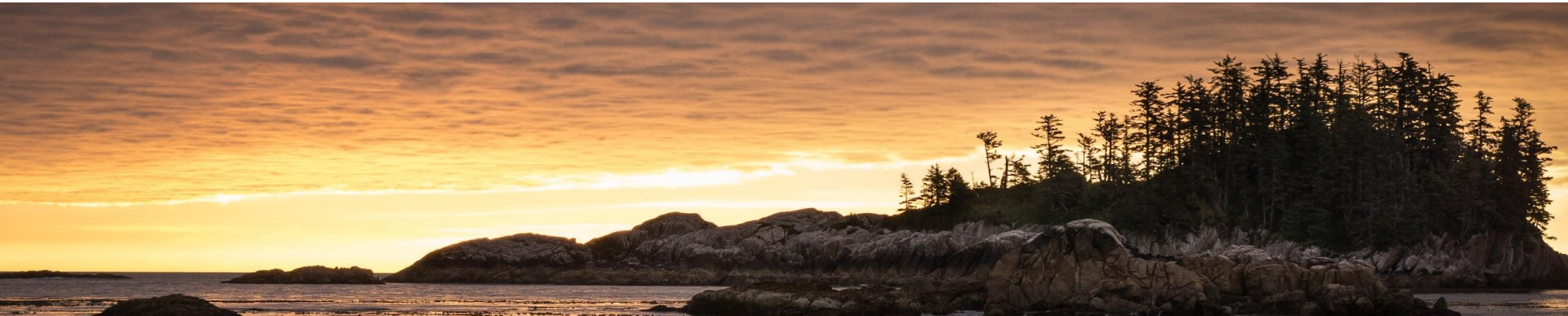
Input Data Formats

- Open source formats where possible
  - CSV, tabular text files, JSON
  - NetCDF files
- Excel data requires format transformation
- Real-time data
  - Requires API or public URL for regular access
  - All data transformations will need to be completely automated

Full list of ERDDAP-supported input file types:
https://coastwatch.pfeg.noaa.gov/erddap/download/setupDatasetsXml.html#EDDTa

Standardizing Dates and Coordinates

- Use ISO 8601 dates
  - YYYY-MM-DD
  - 2022-03-09
- Use ISO 8601 times
  - 2022-03-09T12:00:00-00:00 or 2022-03-09T12:00:00Z
- Always record the time zone that was used
- Ideally use UTC time to make time comparisons easier
- Use decimal degrees for latitude and longitude



Climate and Forecast Conventions

> The conventions define metadata that provide a definitive description of what the data in each variable represents, and the spatial and temporal properties of the data. This enables users of data from different sources to **decide which quantities are comparable**, and facilitates building applications with powerful extraction, regridding, and display capabilities. The CF convention **includes a standard name table**, which defines strings that identify physical quantities.

cfconventions.org

Standardizing Data Variables

- Table setup
  - ERDDAP requires one column for each unique variable
- Use Climate and Forecast (CF) standard variable names where possible
  - The CF standard names table is the standard for non-biological variables
  - Many CF standard names exist to capture exactly what you are measuring
  - CF standard name examples (hundreds to choose from)
    - sea_water_temperature
    - sea_water_salinity
    - eastward_sea_water_velocity
  - Use search to identify the best one to use
- If no CF standard name exists for a variable
  - Can set a custom variable name of your choice
  - Using ‘variable_names_with_underscores’ format can help with consistency
  - Consider requesting a new CF variable via GitHub
- Consider whether all variables need to be published

Standardizing Variable Units

- All variables should have an applicable unit defined in the column heading
  - E.g. sea_water_pressure (dbar)
  - The unit specified must apply to all values in that column
- Climate and Forecast (CF) standard names specify the canonical units
  - E.g. for sea_water_temperature the canonical units are in Kelvin
  - In practice can use any unit (e.g. Celsius) that is convertible to the canonical unit
- CF uses UDUNITS software to provide
  - Units and definitions
  - Standard conversions between unit types

Climate and Forecast Best Practices

- Detailed set of Climate and Forecast standard variables requirements are available
- Variable attributes, dimensions not covered
  - Discrete sampling geometries
- Compliance checker tools exist for ERDDAP/NetCDF
  - CIOOS Atlantic aiming for baseline CF 1.6 compliance for all ERDDAP datasets

Example data transformation

Before

Platform	Date	Time	Lat	Lon	Temp	Pressure	Calibration coeff
Buoy123	2022-03-08	12:00:00	45° 7' 48"	-59° 13' 48"	10	40.2	5.4
…							

After

platform	date	latitude	longitude	sea_water_temperature (Celsius)	sea_water_pressure (dbar)
Buoy123	2022-03-08T16:00:00Z	45.13	59.23	10	40.2
…					

Marine Ecology Lab Example

- Intertidal temperature and tide height data from St. FX Marine Ecology Lab
- Original data made available through Excel file in Figshare here
- Metadata entry form completed
- Transformation script at CIOOS Atlantic
  - Conversion to CSV
  - Renaming columns to CF standards
  - Date and time conversion to UTC in ISO 8601 format
  - Compliance checking
- Final output CIOOS Atlantic catalogue entry with ERDDAP resource link here

FORCE ADCP Example

Waves	Data/Waves	—
BandDirection_L_DirMean	BandDirection L DirMean	1D
BandDirection_L_DirTp	BandDirection L DirTp	1D
BandDirection_L_SprTp	BandDirection L SprTp	1D
BandDirection_U_DirMean	BandDirection U DirMean	1D
BandDirection_U_DirTp	BandDirection U DirTp	1D
BandDirection_U_SprTp	BandDirection U SprTp	1D
BandFrequency_L_High	BandFrequency L High	1D
BandFrequency_L_Low	BandFrequency L Low	1D
BandFrequency_U_High	BandFrequency U High	1D
BandFrequency_U_Low	BandFrequency U Low	1D
BandHeight_L_Hm0	BandHeight L Hm0	1D
BandHeight_U_Hm0	BandHeight U Hm0	1D
BandPeriod_L_Tm02	BandPeriod L Tm02	1D
BandPeriod_L_Tp	BandPeriod L Tp	1D
BandPeriod_U_Tm02	BandPeriod U Tm02	1D
BandPeriod_U_Tp	BandPeriod U Tp	1D
BandSpectrumType_Lower	BandSpectrumType Lower	1D
BandSpectrumType_Upper	BandSpectrumType Upper	1D
Battery_Voltage	Battery Voltage	1D
CurrentDirection	CurrentDirection	1D
CurrentSpeed	CurrentSpeed	1D
Direction	Direction	2D
Direction_DirTp	Direction DirTp	1D
Direction_MeanDir	Direction MeanDir	1D
Direction_SprTp	Direction SprTp	1D
DirectionalSpectra_Frequency	DirectionalSpectra Freque...	1D
DirectionalSpectra_Spread	DirectionalSpectra Spread	2D
EnergySpectra	EnergySpectra	2D
EnergySpectra_Frequency	EnergySpectra Frequency	1D
FourierCoefficients_A1	FourierCoefficients A1	2D
FourierCoefficients_A2	FourierCoefficients A2	2D
FourierCoefficients_B1	FourierCoefficients B1	2D
FourierCoefficients_B2	FourierCoefficients B2	2D

- Final CIOOS Atlantic record here

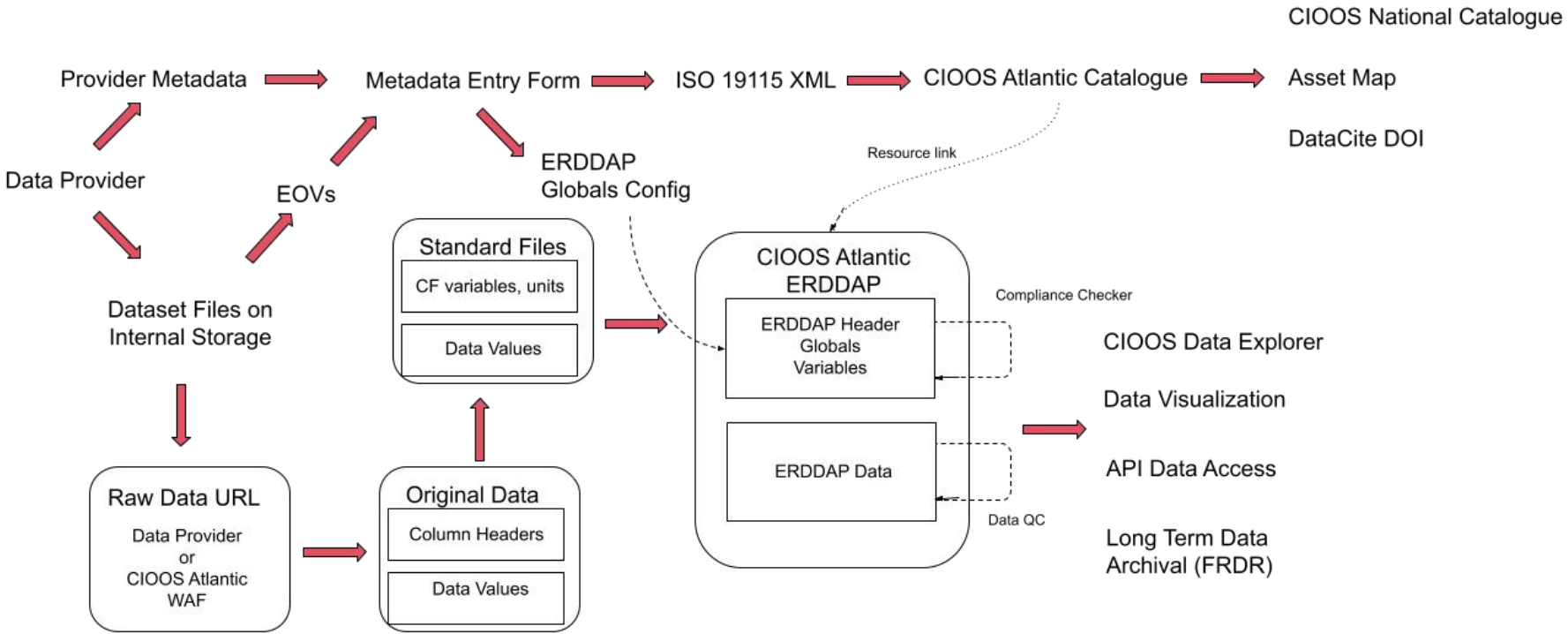
Dataset Title: **FORCE Acoustic Doppler Current Profiler (ADCP)**

Institution: FORCE (Dataset ID: FORCE_Mar2018_ADCP_Waves)

Information: Summary | License | Metadata | Background | Make a graph

Variable	Check All	Uncheck All	Optional Constraint #1	
sea_water_pressure (dbar)			>=	
time (UTC)			>=	2018-05-17T00:00:00Z
sea_surface_wave_maximum_period (s)			>=	
sea_surface_wave_significant_height (m)			>=	

Data Integration Workflow



Next Steps

- CIOOS Data Explorer (CDE)
  - All standardized data in ERDDAP can be viewed/downloaded in CDE
- Persistent identifiers and digital object identifiers (DOIs)
  - CIOOS Atlantic can mint DOIs for our catalogue records using DataCite Canada Consortium (DCAN)
- Federated Research Data Repository (FRDR) for long term archival
  - Static files can be archived for > 10 years with FRDR
- Application Programming Interface (API) data access
  - Programmatic access to data (erddapy) and metadata (CKAN API)
  - Allow your queries and analyses to be automatically reprocessed as new data is made available
  - Example of Jupyter erddapy data access and visualization here

Questions?

CIOOS ATLANTIC

REGIONAL ASSOCIATION OF THE
CANADIAN INTEGRATED OCEAN OBSERVING SYSTEM

This project would not be possible without the financial support and the continued involvement of these organizations :

Fisheries and Oceans Canada
Pêches et Océans Canada

Tula
TULA FOUNDATION

MEOPAR

cioosatlantic.ca| cioos.ca

info@cioosatlantic.ca

@cioos.siooc

@CIOOSAtlantic