

SECOORA DATA PORTAL
Centralized access to Southeast U.S. coastal and ocean data

EXPLORE REAL TIME DATA | SEARCH 2300+ DATASETS | GLIDERS

SECOORA Data Portal
Go to version: 1.0 portal »



Welcome to the new SECOORA Data Portal! Use the portal to:

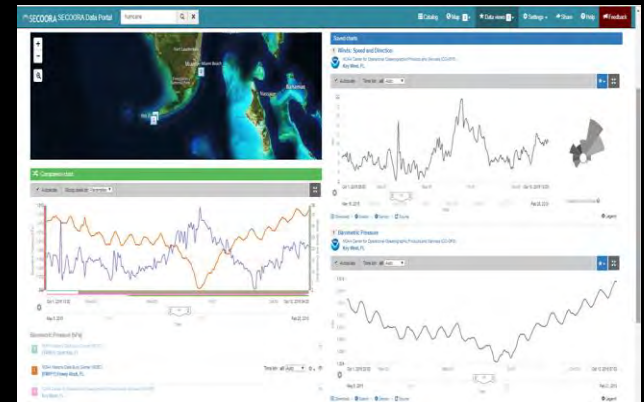
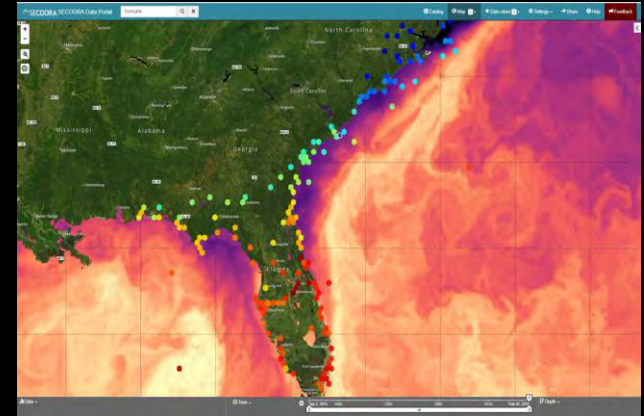
- Search and download real-time data
- Search historical data
- Compare datasets from different stations
- Generate and share custom data views (how to coming soon!)
- Access metadata for SECOORA stations
- Access to regional and sub-regional models, including coastal circulation, water quality and fisheries habitat models.

Coming soon the portal will provide access to regional and sub-regional models, including coastal circulation, water quality and fisheries habitat models.

Explore map | Catalog | Glider deployments

Release notes | Documentation

For best results, use the latest version of these browsers:  



Rob Bochenek - Axiom Data Science

FROM THE IOOS COMMUNITY: WE MISS YOU



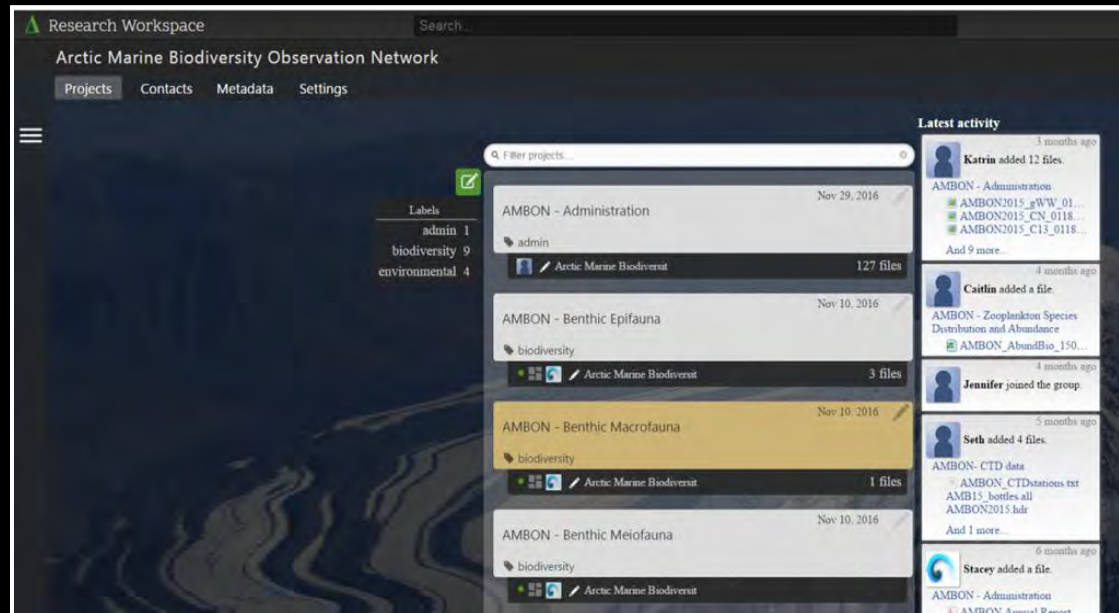
The National Science Foundation defines cyberinfrastructure as:

In scientific usage, cyberinfrastructure is a technological and sociological solution to the problem of efficiently connecting laboratories, data, computers, and people with the goal of enabling derivation of novel scientific theories and knowledge.

- Understanding of the existing community developed data standards, protocols and software
- Scalable compute and storage infrastructure (HPC)
- Human capacity - data scientists, data librarians, data coordinators, software engineers...
- Science community that can benefit from support

RESEARCH WORKSPACE

- Organize into projects, research campaigns and organizations
- Coordinate data exchange across networks, groups, programs
- ISO 19110/19115-2 standards metadata editor
- Execute server side R and Python numeric workflows (Jupyter) on uploaded data AND any data in SECOORA DMAC System
- Archive pathway to DataONE & Datacite DOI minting



LOOKING TO THE FUTURE OF SECOORA DMAC



- Now that the infrastructure is built...it's time to populate it with more SECOORA data
- Interface SECOORA community with leveraged programs and capabilities (ATN, MBON, NOAA HPCC, COMT, OTT)
- More direct science community support

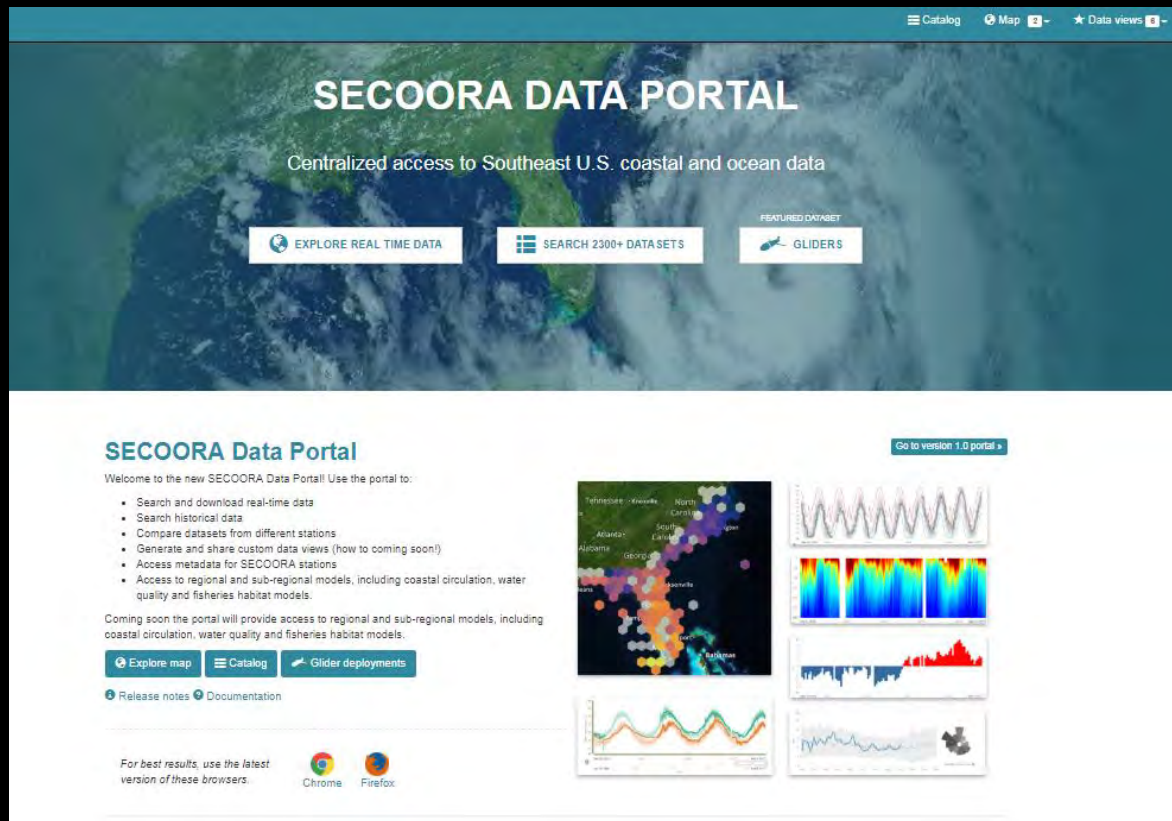


Stacey Buckelew - Axiom Data Science

SESSION OVERVIEW

- Portal demo
- Guided example
- Self-paced exercise(s)

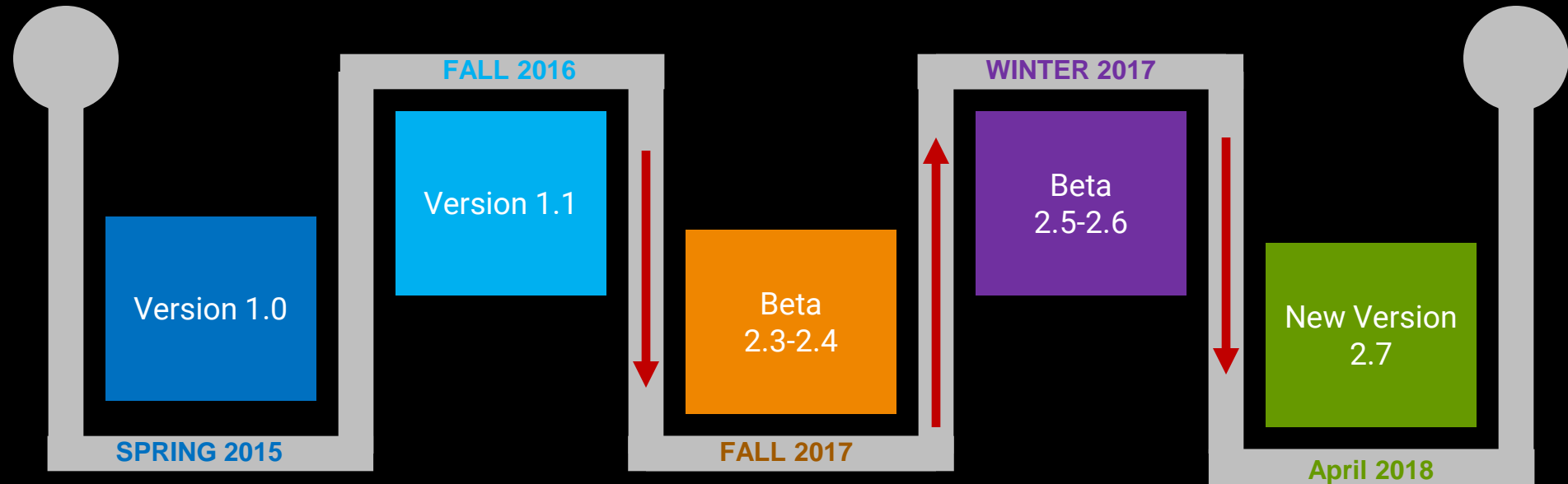




The screenshot shows the SECOORA Data Portal v2.7 homepage. The header features a navigation bar with 'Catalog', 'Map', and 'Data views' options. The main banner displays the title 'SECOORA DATA PORTAL' and the subtitle 'Centralized access to Southeast U.S. coastal and ocean data'. Below the banner are three prominent buttons: 'EXPLORE REAL TIME DATA', 'SEARCH 2300+ DATA SETS', and 'FEATURED DATASET GLIDERS'. The main content area includes a 'SECOORA Data Portal' section with a welcome message and a list of capabilities: searching and downloading real-time data, searching historical data, comparing datasets from different stations, generating and sharing custom data views, accessing metadata for SECOORA stations, and accessing regional and sub-regional models. A 'Go to version 1.0 portal' link is also present. Below the text are buttons for 'Explore map', 'Catalog', and 'Glider deployments', along with links for 'Release notes' and 'Documentation'. At the bottom, there is a note about browser compatibility with Chrome and Firefox icons.

Customizable data environment allowing scientists, managers, and the public others to discover, access, and interact with coastal ocean data from different sources

BACKGROUND



- Feedback from science user community
- Community standards, protocols and software
- Scalable compute and storage infrastructure (HPC)

NEED

Remote data exchange to support analysis, integration, & collaboration



CAPABILITY

- Data discovery and access
- Data visualization
- Dynamic interaction
- Custom exploration
- State saving & sharing
- Data publication & archive



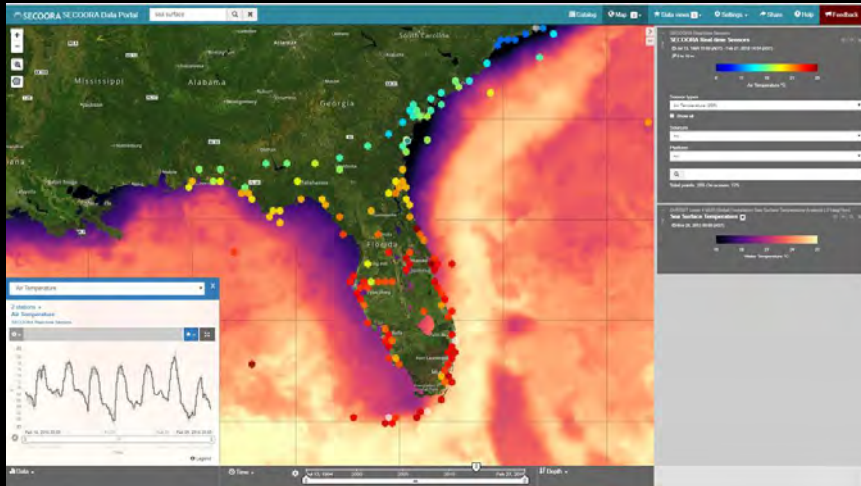
AMOUNT

- 2,300 data layers
- 1,500 sensors
- 35 parameters
- 20+ data sources
- 5 million obs/ week



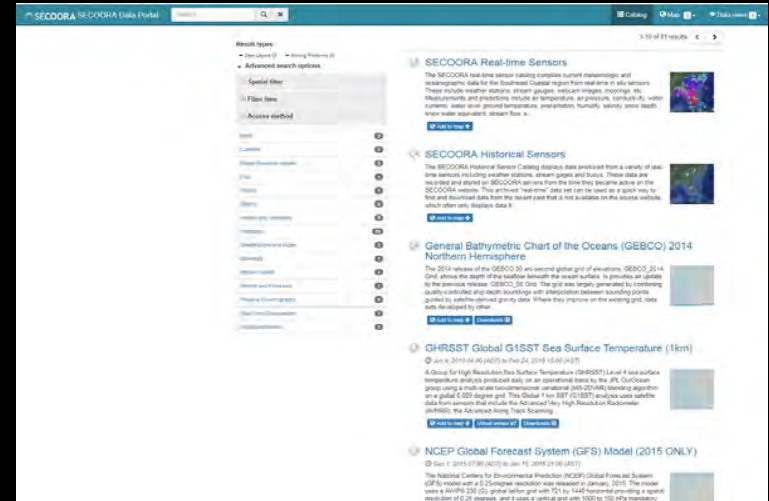
Map

Integrate & visualize data from many sources



Catalog

Search, metadata, & data download



Data Views

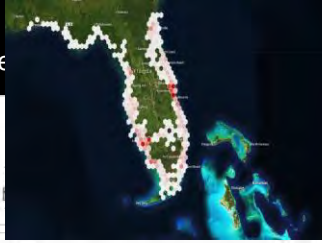
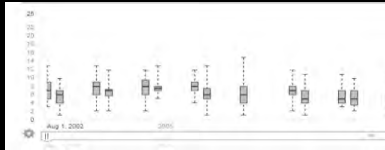
Rapidly assimilate & compare different data streams



Data Types

Biodiversity

count, richness, diversity indices



Platforms

moorings, shore stations



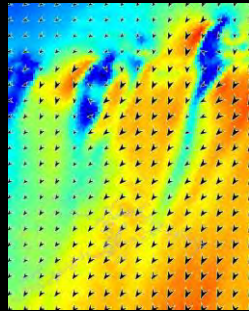
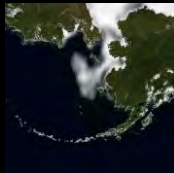
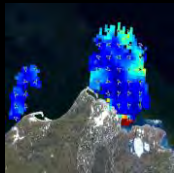
Products

skill assessment, shoreline change, etc.



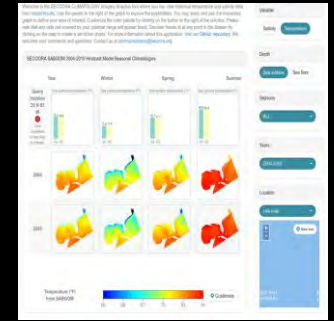
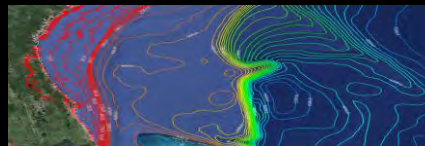
Grids

models, satellite, radar

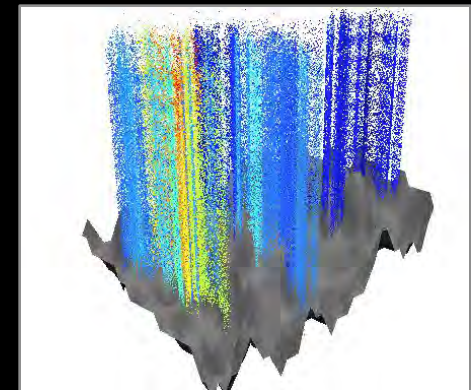
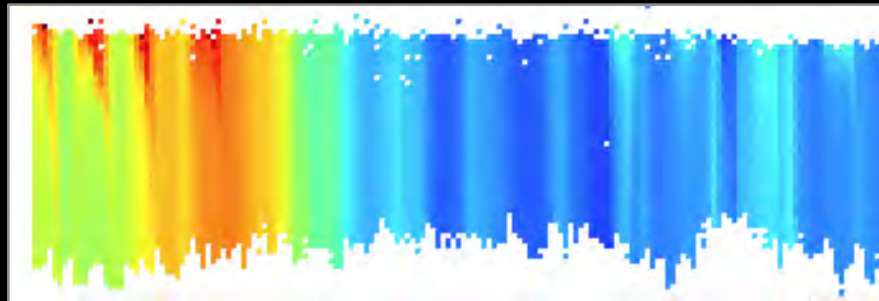


GIS

Habitat types, bathymetry, fishing zones, etc.




Gliders



Live Demo

<https://portal.secoora.org/>



The screenshot shows the SECOORA Data Portal homepage. The header features a teal navigation bar with links for Catalog, Map, and Data views. The main banner displays the title "SECOORA DATA PORTAL" and the subtitle "Centralized access to Southeast U.S. coastal and ocean data". Below the banner are three buttons: "EXPLORE REAL TIME DATA", "SEARCH 2300+ DATA SETS", and "GLIDER'S". The main content area includes a "SECOORA Data Portal" section with a welcome message and a list of features. To the right, there is a "Go to version 1.0 portal" link and a grid of data visualizations, including a map of the Southeast U.S. and several time-series plots. At the bottom, there are links for "Explore map", "Catalog", and "Glider deployments", along with "Release notes" and "Documentation". A note at the bottom left recommends using the latest version of Chrome or Firefox browsers.

SECOORA DATA PORTAL

Centralized access to Southeast U.S. coastal and ocean data

EXPLORE REAL TIME DATA SEARCH 2300+ DATA SETS GLIDER'S

Go to version 1.0 portal >

SECOORA Data Portal

Welcome to the new SECOORA Data Portal! Use the portal to:

- Search and download real-time data
- Search historical data
- Compare datasets from different stations
- Generate and share custom data views (how to coming soon!)
- Access metadata for SECOORA stations
- Access to regional and sub-regional models, including coastal circulation, water quality and fisheries habitat models.

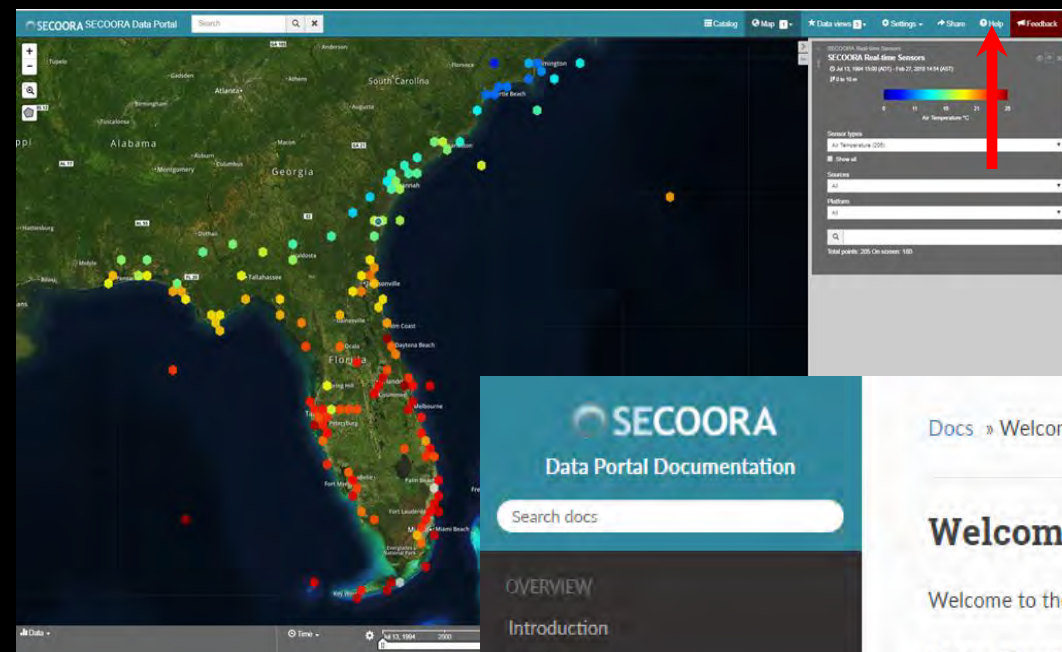
Coming soon the portal will provide access to regional and sub-regional models, including coastal circulation, water quality and fisheries habitat models.

Explore map Catalog Glider deployments

Release notes Documentation

For best results, use the latest version of these browsers:

Chrome Firefox



SECOORA
Data Portal Documentation

Search docs

OVERVIEW

- Introduction
- Catalog
- Data Views

HOW-TO

- Catalog
- Map

[Docs](#) » [Welcome](#)

[View page source](#)

Welcome

Welcome to the help site for the SECOORA Data Portal!

Data Portal Overview

The Data Portal is a data exploration tool with a customized public web interface that allows scientists, managers, and the general public to discover and access public data.

The Data Portal integrates datasets from many different sources. You can search or browse real-time conditions, operational and research forecasts, satellite observations, and other spatially referenced datasets that describe the regional biological, chemical, and physical characteristics.

Datasets in the Data Portal can be interactively mapped or charted using advanced features, such as the ability to create comparisons between data sources, bin data by time, and plot climatologies and anomalies.

Additionally, you can create and share custom compilations of sensor and model outputs to spotlight environmental events or geographic locations, access metadata and project contacts, and download datasets in a variety of formats.

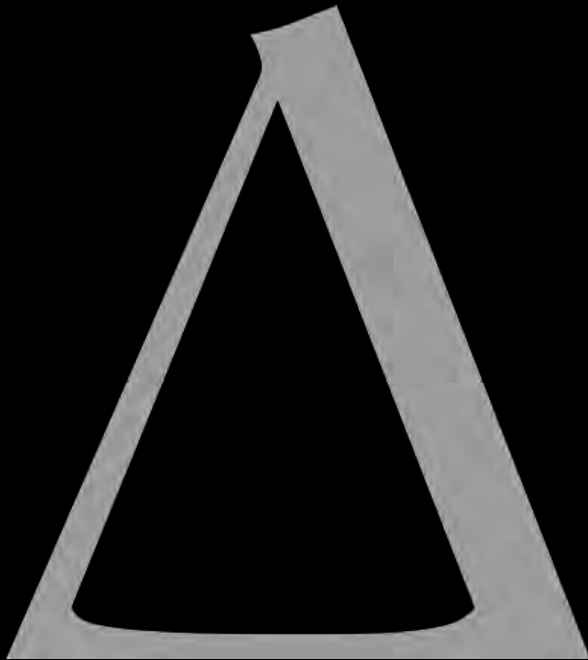
The Data Portal is being actively developed by [Axiom Data Science](https://axiomdatascience.com) and is currently in Beta status. Several new features are scheduled for release in 2018, as indicated in the text. For the notes about the latest portal version please visit: <https://axiomdatascience.com/portal-updates/>

Documentation Overview

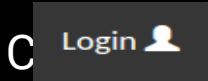

The Break: Checklist

1. Get Coffee
2. Log in to Research Workspace
3. Clone project (if you want to follow along)

The Research Workspace

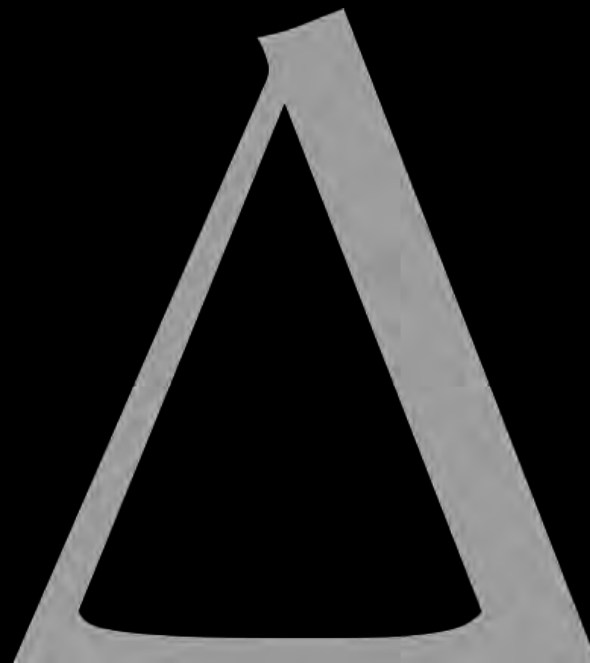


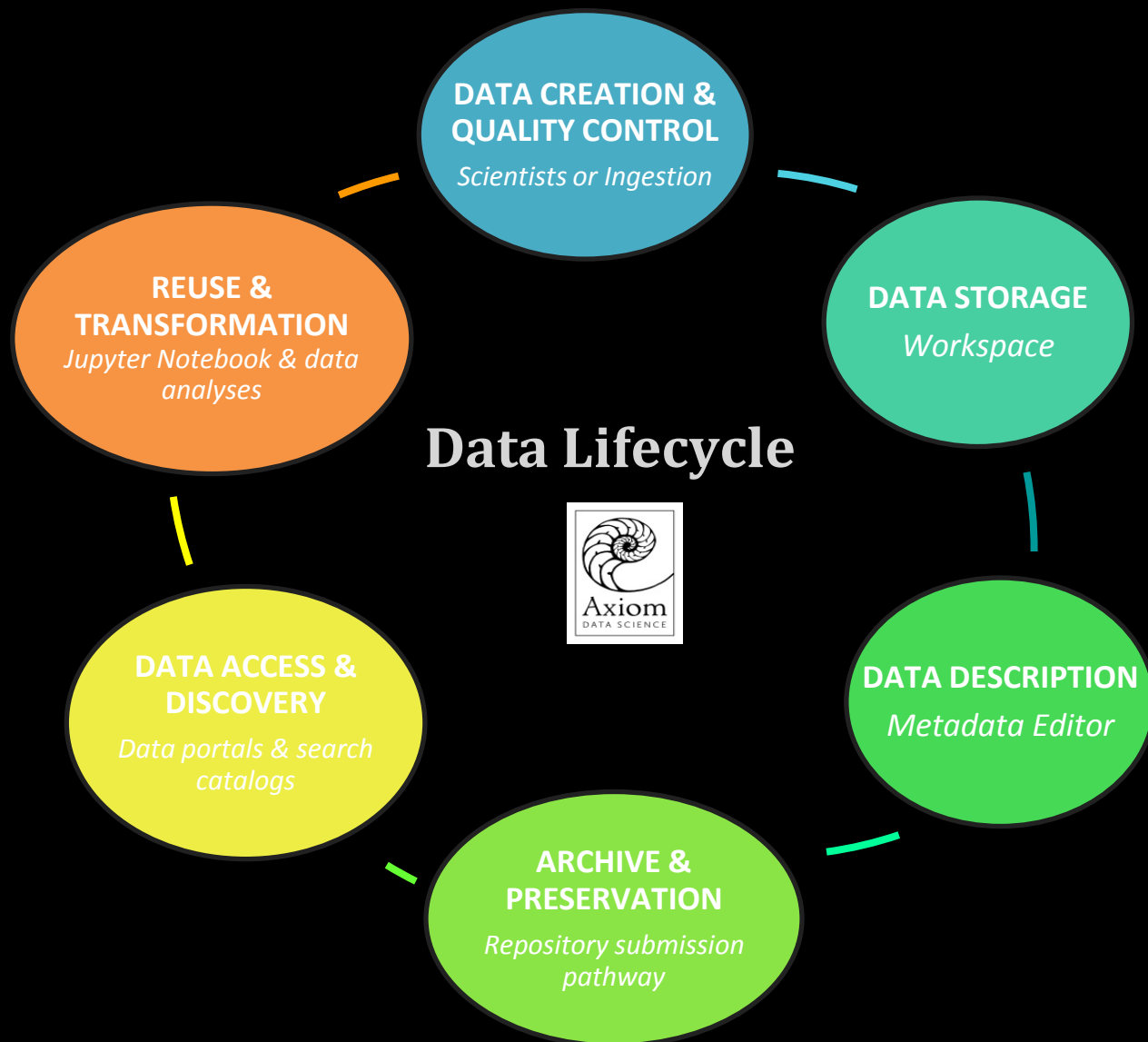
- Logging in

1. <https://researchworkspace.com>
2.  button
3. Click 'Create Account'.
4. Access code: workspace2018
5. Enter email address and password (or click the  button if your email is Google-authorized).

Session Overview

- Getting Started in the Research Workspace
- Clone a Project
- Jupyter Notebooks
 - *practice....*
- Feedback





DATA CREATION & QUALITY CONTROL

Scientists or Ingestion

DATA STORAGE

Workspace

DATA DESCRIPTION

Metadata Editor

ARCHIVE & PRESERVATION

Repository submission pathway

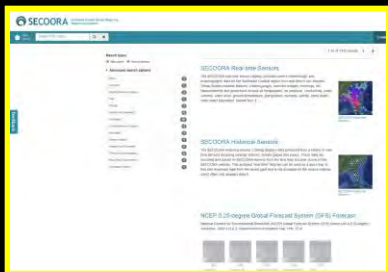
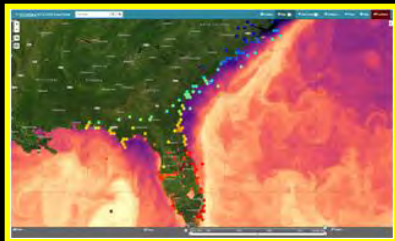
REUSE & TRANSFORMATION

Jupyter Notebook & data analyses

DATA ACCESS & DISCOVERY

Data portals & search catalogs

Data Lifecycle



DATA CREATION & QUALITY CONTROL

Scientists or Ingestion

DATA STORAGE

Workspace

DATA DESCRIPTION

Metadata Editor

ARCHIVE & PRESERVATION

Repository submission pathway

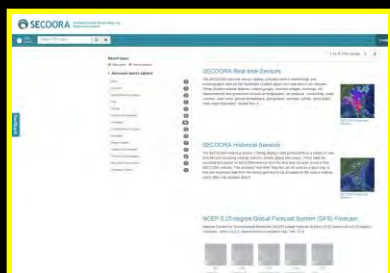
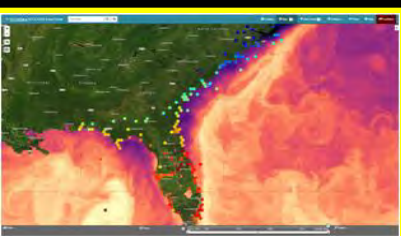
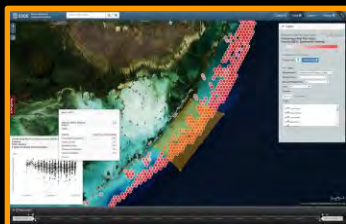
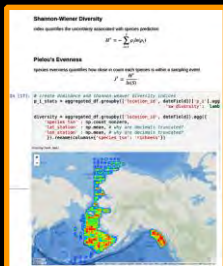
REUSE & TRANSFORMATION

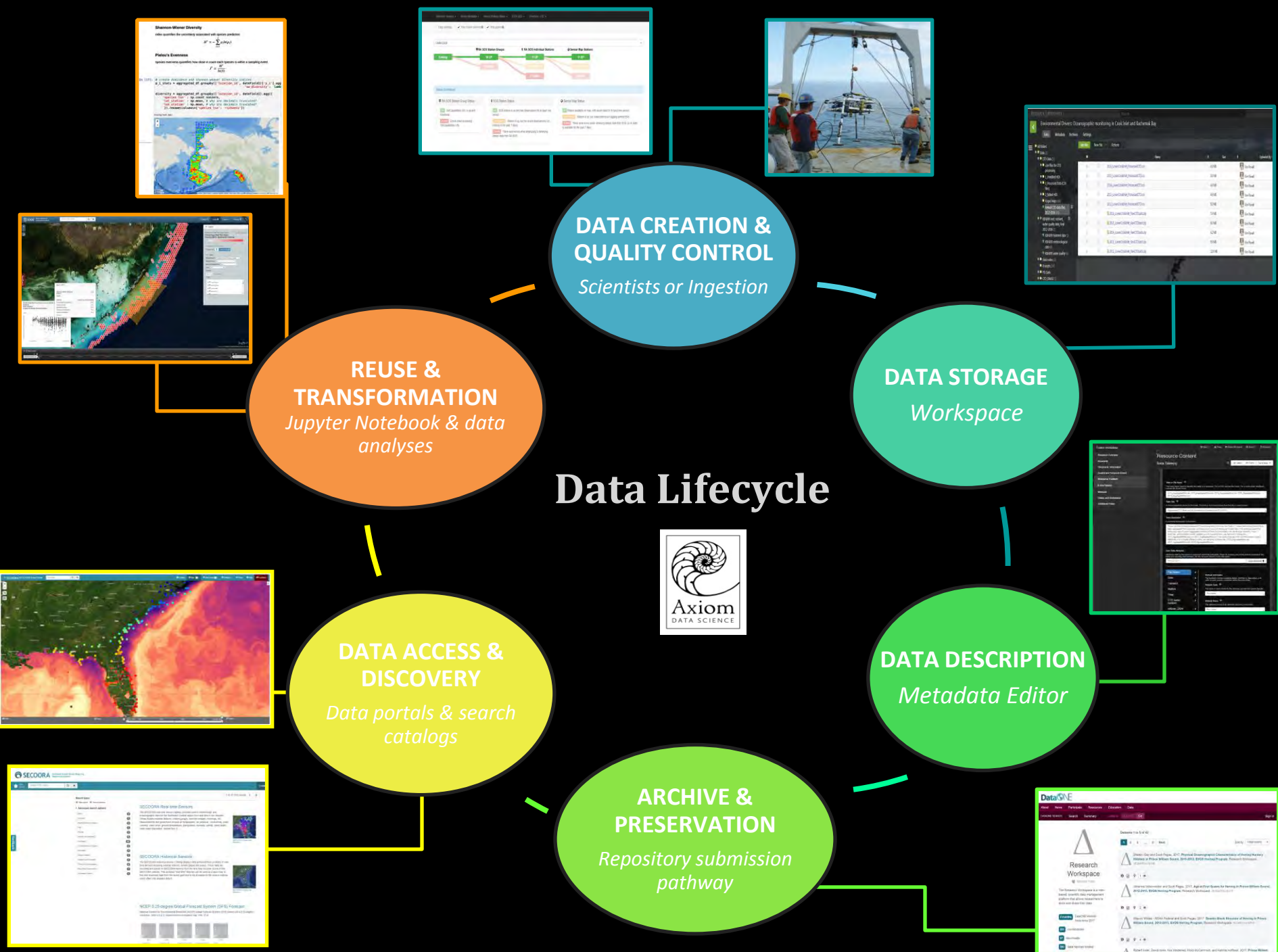
Jupyter Notebook & data analyses

DATA ACCESS & DISCOVERY

Data portals & search catalogs

Data Lifecycle

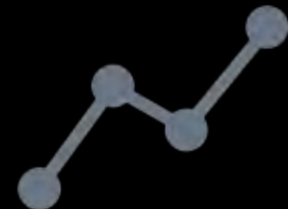




RESEARCH Δ WORKSPACE

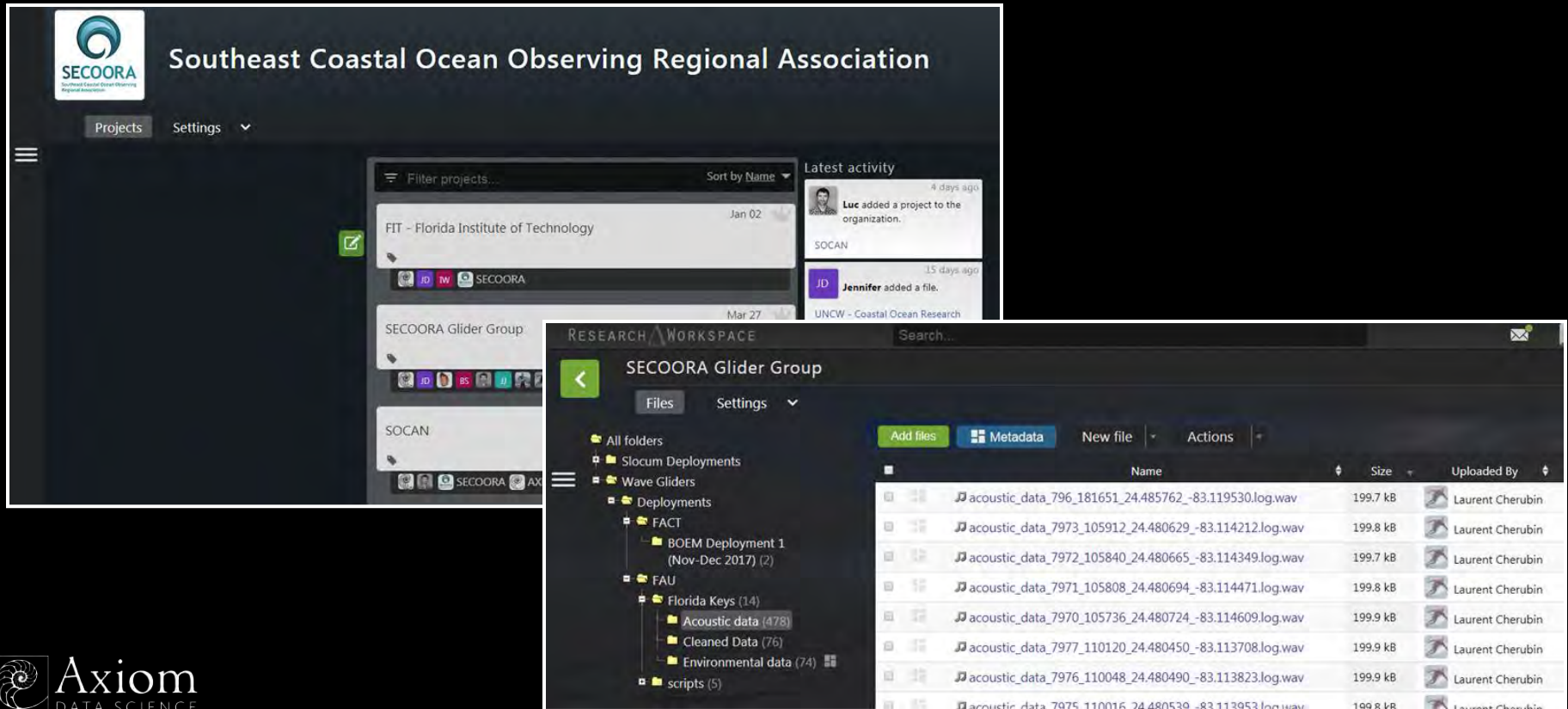
~web-based platform for collaboratively managing science projects through the entire data lifecycle~

Share
Analyze
Preserve



RESEARCH WORKSPACE: Store & Share

- Organize data into projects, research campaigns & organizations for coordinated data exchange
- Manage sharing through advanced security permissions



The screenshot displays the Research Workspace interface for the Southeast Coastal Ocean Observing Regional Association (SECOORA). The main view shows the 'SECOORA Glider Group' project, which is organized into folders: 'All folders', 'Slocum Deployments', 'Wave Gliders', 'Deployments', 'FACT', 'BOEM Deployment 1 (Nov-Dec 2017) (2)', 'FAU', 'Florida Keys (14)', 'Acoustic data (478)', 'Cleaned Data (76)', 'Environmental data (74)', and 'scripts (5)'. A table of acoustic data files is shown, listing the file name, size, and the user who uploaded it.

Name	Size	Uploaded By
acoustic_data_796_181651_24.485762_-83.119530.log.wav	199.7 kB	Laurent Cherubin
acoustic_data_7973_105912_24.480629_-83.114212.log.wav	199.8 kB	Laurent Cherubin
acoustic_data_7972_105840_24.480665_-83.114349.log.wav	199.7 kB	Laurent Cherubin
acoustic_data_7971_105808_24.480694_-83.114471.log.wav	199.8 kB	Laurent Cherubin
acoustic_data_7970_105736_24.480724_-83.114609.log.wav	199.9 kB	Laurent Cherubin
acoustic_data_7977_110120_24.480450_-83.113708.log.wav	199.9 kB	Laurent Cherubin
acoustic_data_7976_110048_24.480490_-83.113823.log.wav	199.9 kB	Laurent Cherubin
acoustic_data_7975_110016_24.480530_-83.113953.log.wav	199.8 kB	Laurent Cherubin

- Integrated ISO 19110/19115-2 standards metadata editor
- Author metadata alongside data
- Implement labor-saving tools

Folder metadata

Resource Overview

Basic Overview

Contacts

Category and Form

Keywords

Taxonomic Information

Spatial and Temporal Extent

Resource Content

Methods

Status and Distribution

Additional Fields

Save Copy Delete full record Export Revisions

Resource Overview

Basic Overview Copy Save Next step

This section provides an overview about the dataset and any associated file(s).

Resource Title

A descriptive title for the data that includes the subject matter, where data was collected, and when it was collected.

Assessing abundance of beluga whales in Bristol Bay using genetic mark-recapture methods, 2002-2011

Abstract

A summary of the key aspects of the dataset that includes when, where, why, and how it was collected, as well as a brief description of its variables and file formats.

This project estimated the abundance of beluga whales within the Bristol Bay stock using genetic mark-recapture methods and combined genetic data with aerial survey data to develop an unbiased correction factor for use in future aerial surveys. The project was started in 2004 by the Alaska Beluga Whale Committee, which funded sample collection from 2004 until 2012 and genotyping from 2004 until 2011, and continued through funding from the North Pacific Research Board (NPRB 1516) from 2015 through 2017. The data for this project were generated using genetic markers from skin biopsies of beluga whales Bristol Bay from 2002 to 2011 using mark-recapture methods.

Data from this project consists of 2 .csv data files archived here (NPRB_1516_Bristol_Bay_beluga_whale_abundance_data_sample_list.csv and NPRB_1516_Bristol_Bay_beluga_whale_abundance_data_matching_file.csv).

Purpose

The intention of the dataset and why it was collected or developed, as well as a statement about the dataset's relevance to any larger project or effort.

The Bristol Bay beluga whale stock is genetically distinct from other stocks and tagging studies show it is restricted to Bristol Bay year-round. Quantifying the abundance of belugas in the Bristol Bay stock is important for their management and is critical information for upcoming stock status reviews. This is the first estimate of abundance of belugas in Bristol Bay with appropriate confidence limits.

RESEARCH WORKSPACE Archive

: Publish &

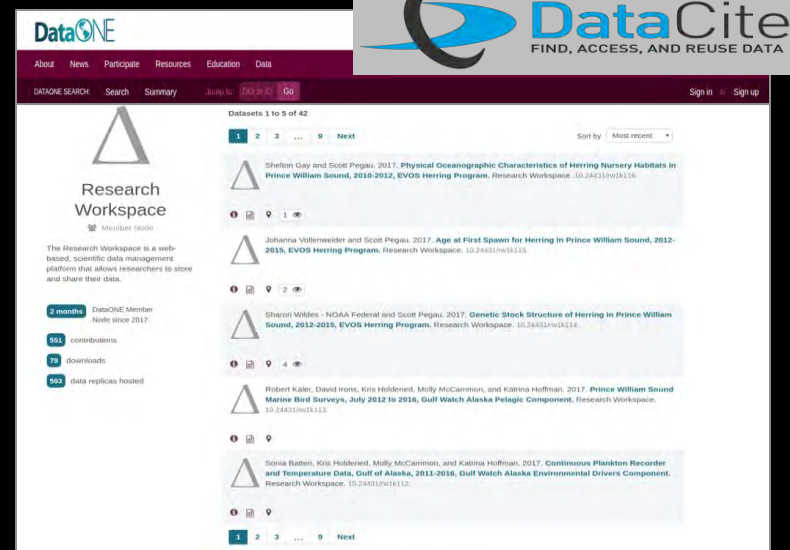
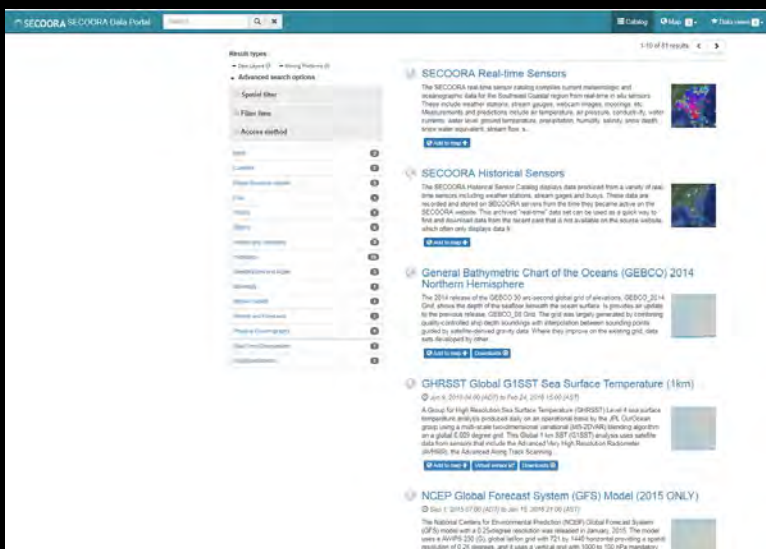
Publish

Archive

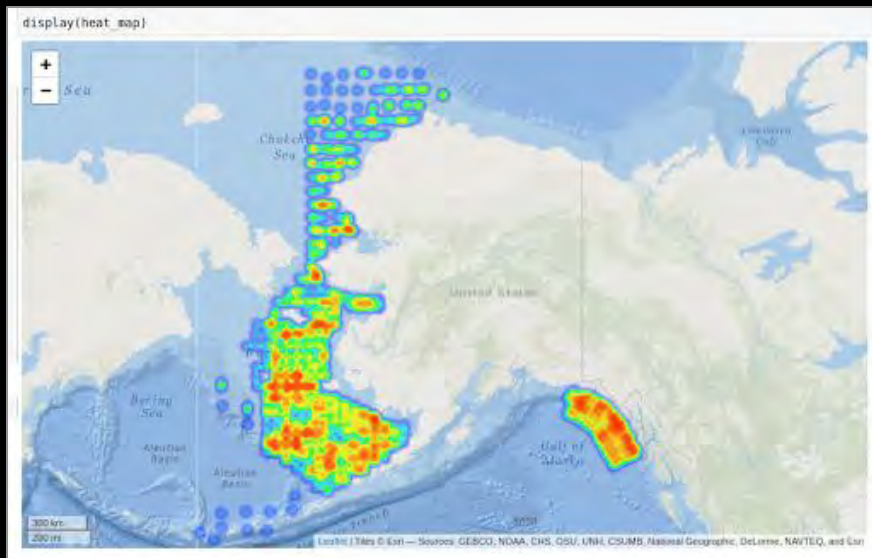
Links



DataONE



- Create and share documents that contain code, equations, and visualizations
- Reproducible numerical simulations and statistical modeling
- Access uploaded data stored in the Workspace or portal



Richness

the number of distinct species found in a sample

$$S = \sum (p_i > 0)$$

% Dominance (Berger-Parker)

the ratio between the number of individuals belonging to the most abundant species and the total number of individuals in the sample

$$\text{Dominance} = \max(p_i)$$

Shannon-Wiener Diversity

index quantifies the uncertainty associated with species prediction

$$H' = - \sum_{i=1}^S p_i \ln(p_i)$$

Pielou's Evenness

species evenness quantifies how close in count each species is within a sampling event

$$J' = \frac{H'}{\ln(S)}$$

```
In [17]: # create dominance and shannon-weaver diversity indices
p_i_stats = aggregated_df.groupby(['location_id', dateField])['p_i'].agg(
    'sw_diversity': lamb

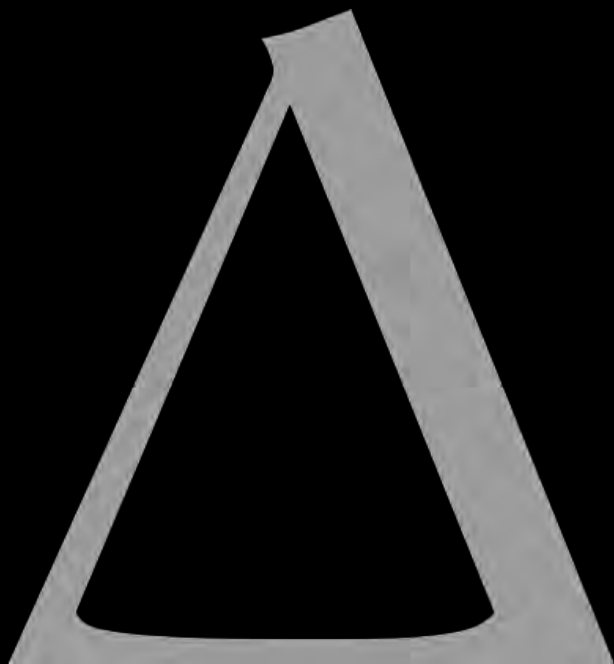
diversity = aggregated_df.groupby(['location_id', dateField]).agg({
    'species_tsn': np.count_nonzero,
    'lat_station': np.mean, # why are decimals truncated?
    'lon_station': np.mean, # why are decimals truncated?
}).rename(columns={'species_tsn': 'richness'})

diversity = diversity.merge(p_i_stats, left_index=True, right_index=True)

# add Pielou's Evenness Index
diversity['evenness'] = diversity['sw_diversity']/np.log(diversity['richness'])

diversity = diversity.reset_index(level=[dateField, 'location_id'])
diversity
```

Research Workspace & Metadata



Resources

Help Docs

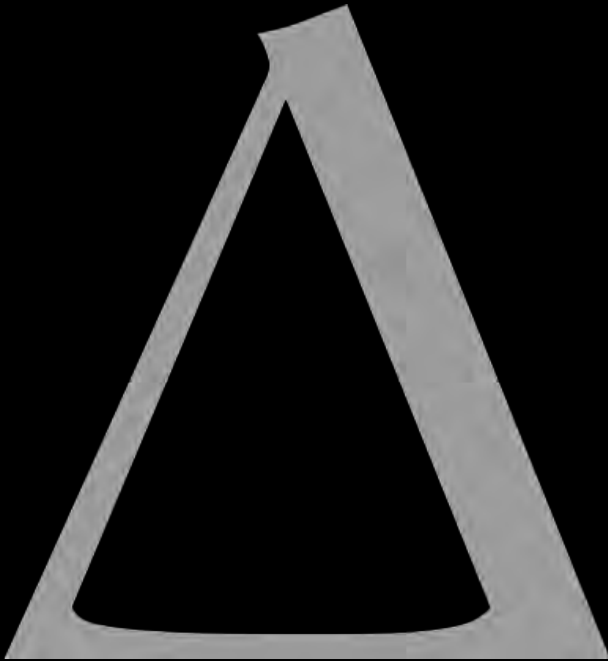
<https://researchworkspace.com/help/>

Technical Assistance

dmac@secoora.org

Research Workspace: Live Demo

<https://researchworkspace.com>



SECOORA Stakeholder Feedback

- What ideas do you have for use applications in your work?
- Are there additional features that would support or improve your science or management throughput?
- Are you using additional data types or feeds that could be made available through SECOORA?

Portal Feedback & Evaluation

<https://goo.gl/forms/bjiFrII5JNy0Lwca>

2

