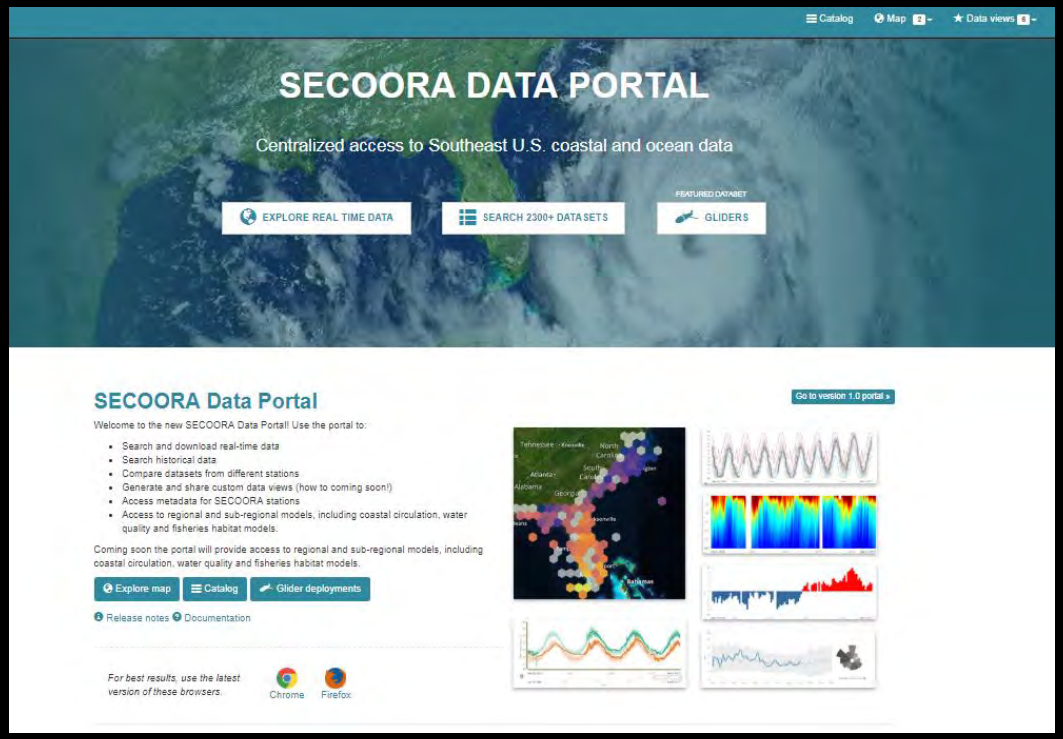
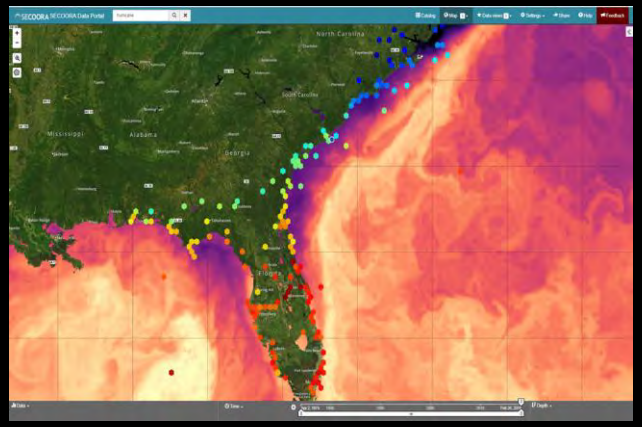
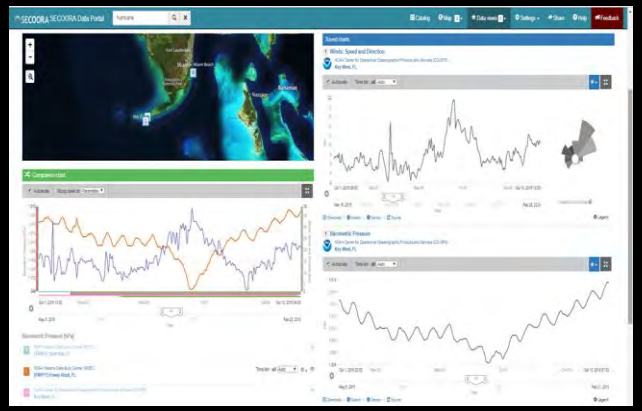
SECOORA DMAC SYSTEM







Rob Bochenek - Axiom Data Science

Axiom DATA SCIENCE

IOOS

From the IOOS Community: We Miss You

SUPPORTING SECOORA THROUGH CYBERINFRASTRUCTU

The National Science Foundation defines cyberinfrastructure as:

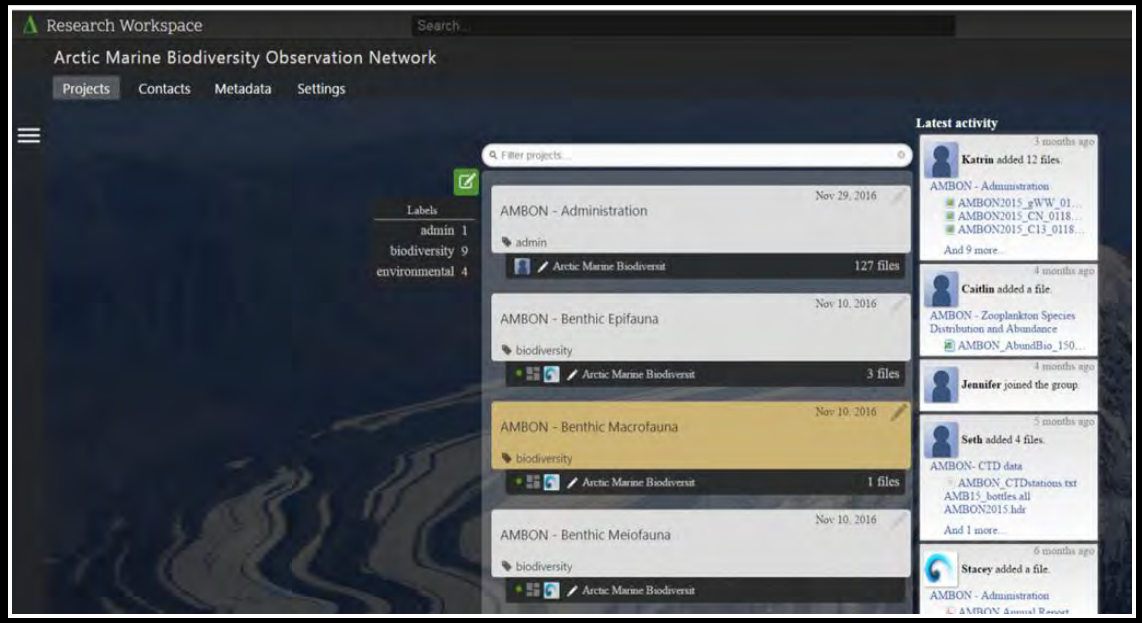
In scientific usage, cyberinfrastructure is a technological and sociological solution to the problem of efficiently connecting laboratories, data, computers, and people with the goal of enabling derivation of novel scientific theories and knowledge.

- 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

Axiom DATA SCIENCE

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



Axiom DATA SCIENCE

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

SECOORA Data Portal

Stacey Buckelew - Axiom Data Science

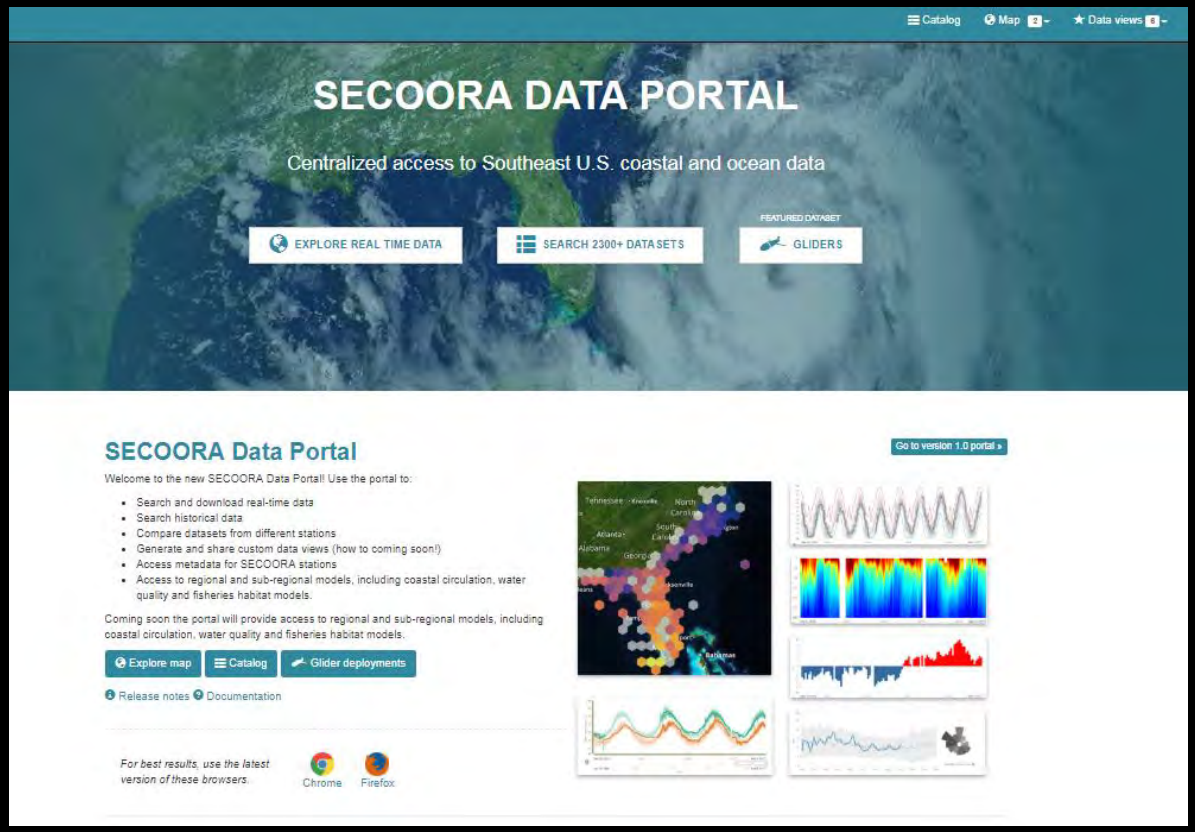
Session Overview

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

Axiom DATA SCIENCE

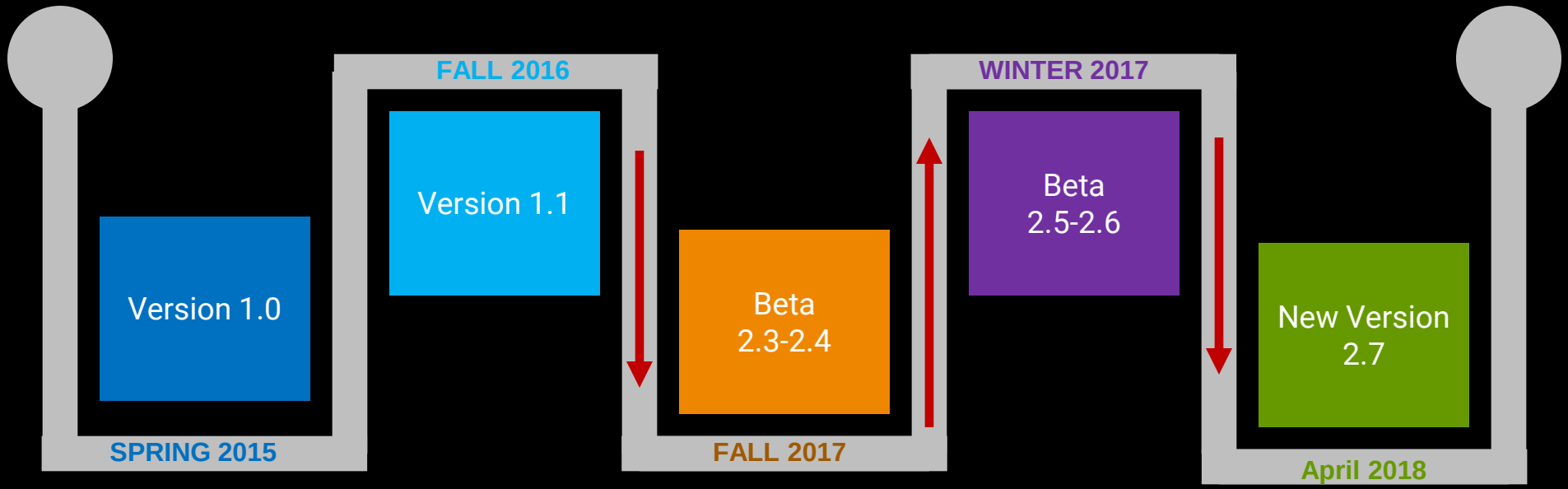
SECOORA Data Portal (v2.7)

https://portal.secoora.org/



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)

Axiom
DATA SCIENCE

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

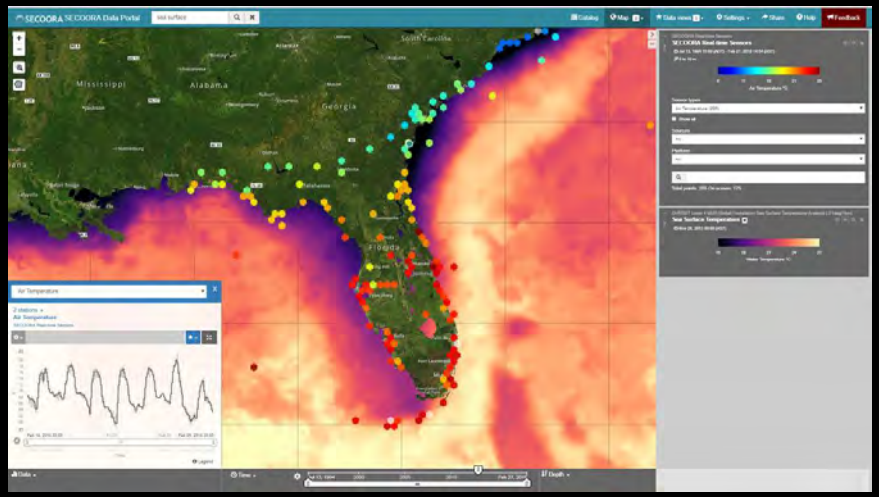
Axiom DATA SCIENCE

SECOORA
Southeast Coastal Ocean Observing Regional Association

https://portal.secoora.org/

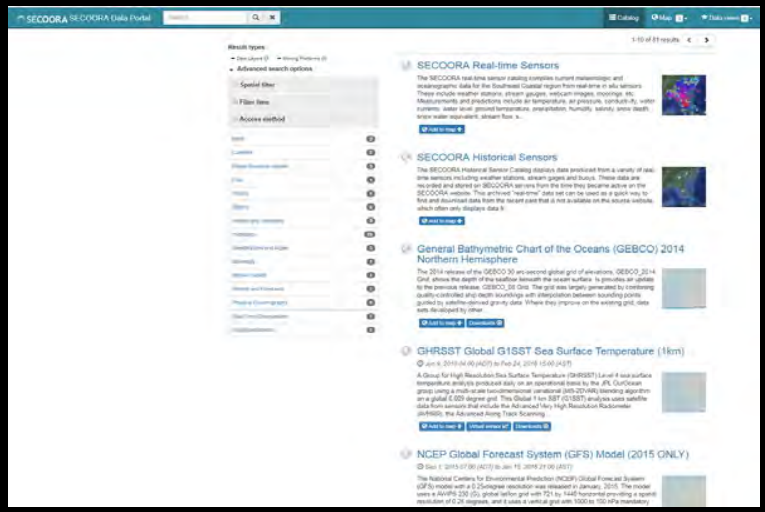
Map

Integrate & visualize data from many sources



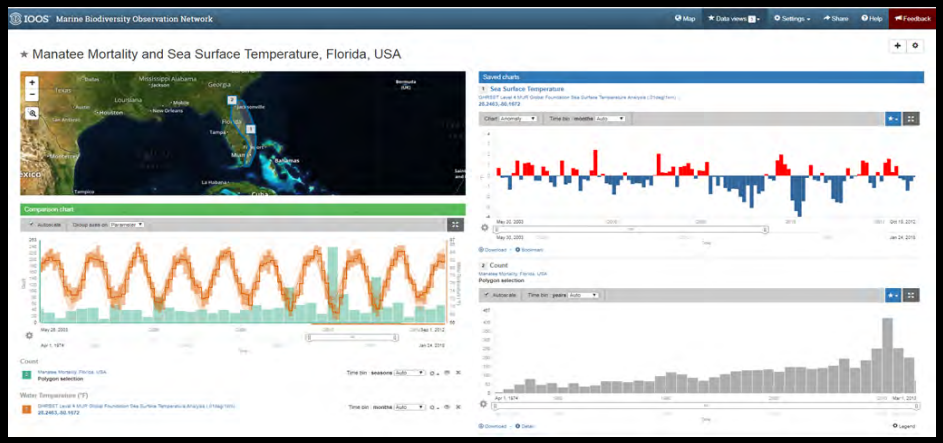
Catalog

Search , metadata, & data download



Data Views

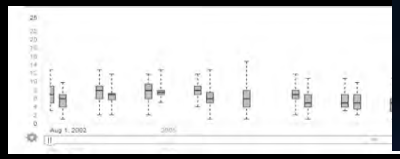
Rapidly assimilate & compare different data streams

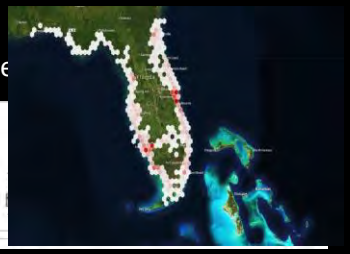


Data Types

Biodiversity

count, richness, diversity indice



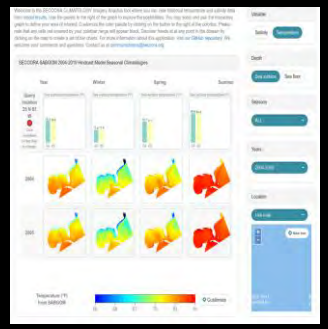


Platforms

moorings, shore stations

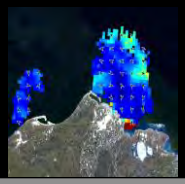
Products

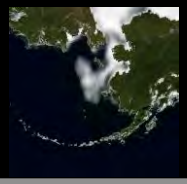
skill assessment, shoreline change, etc.

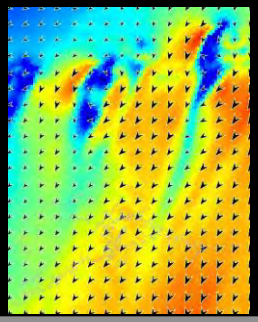


Grids

models, satellite, radar

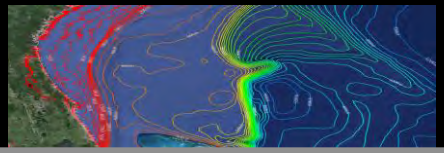




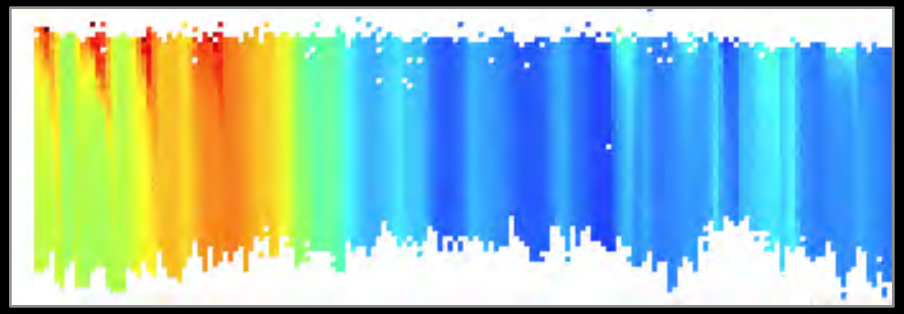


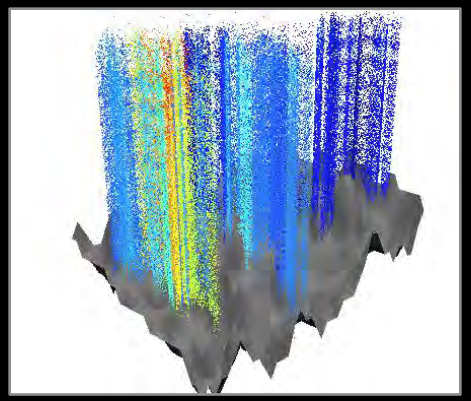
GIS

Habitat types, bathymetry, fishing zones, etc.



Gliders

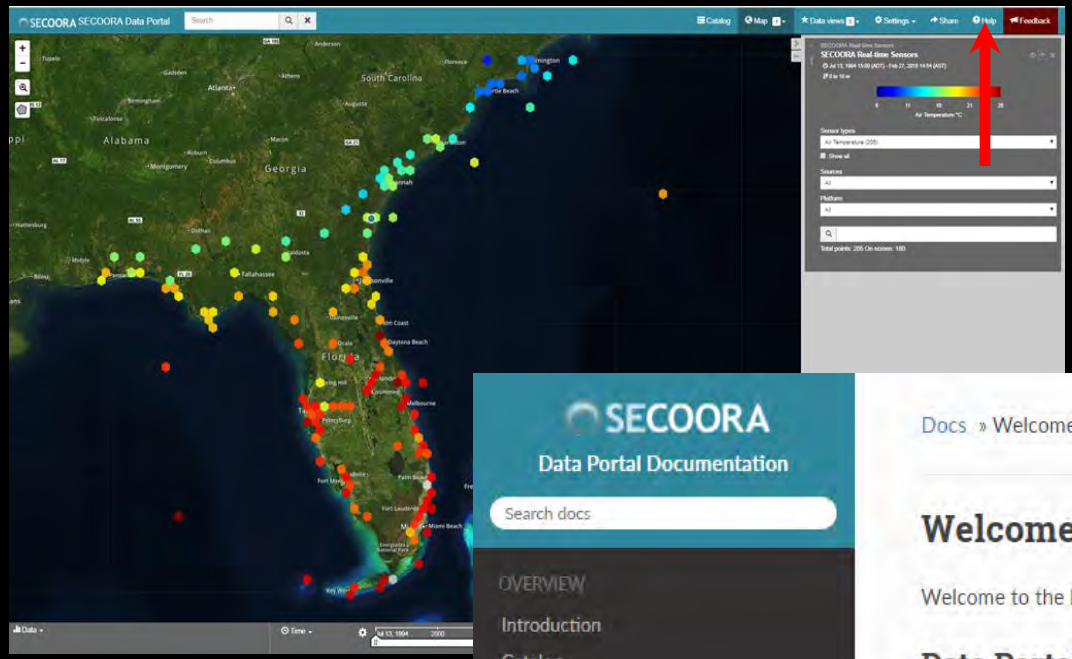




Axiom DATA SCIENCE

Live Demo
https://portal.secoora.org/

Axiom DATA SCIENCE



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 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/

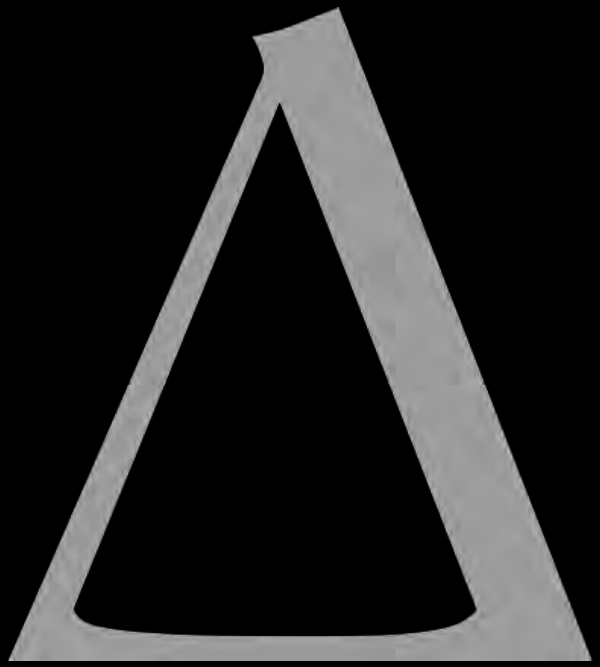
Documentation Overview

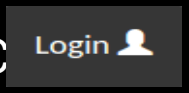
Axiom DATA SCIENCE

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. C 

button
  3. Click ‘Create Account’.
  4. Access code: workspace2018
  5. Enter email address and password (or click the

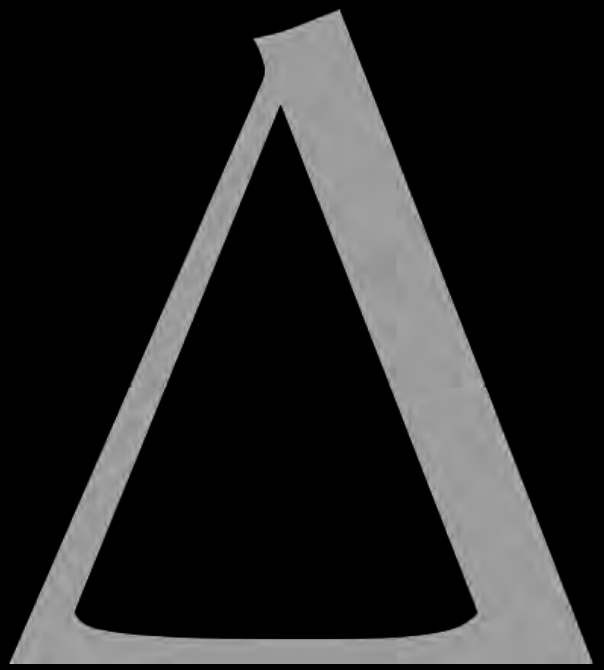
button if your email is Google-authorized).

Axiom DATA SCIENCE

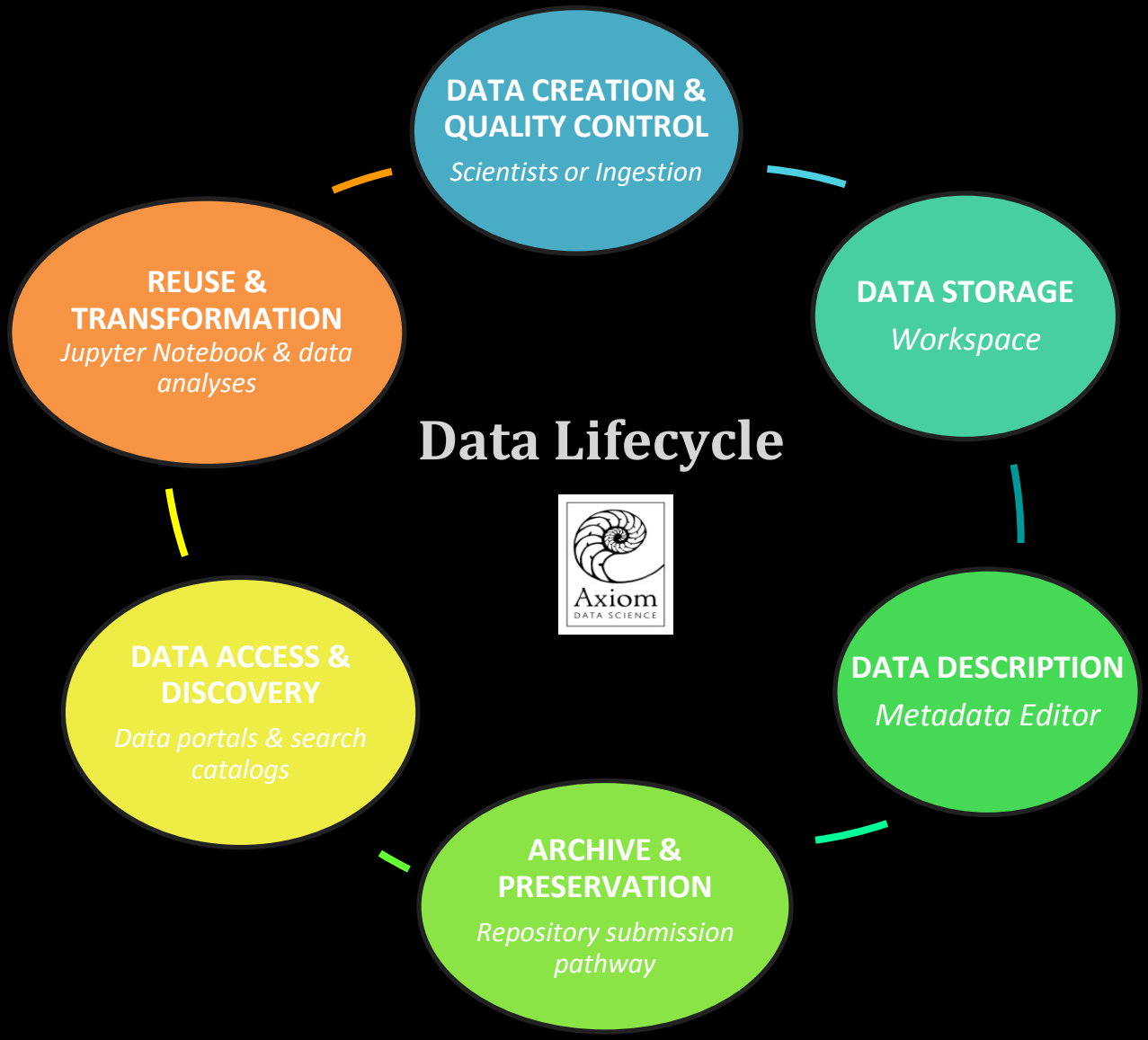
Session Overview

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

- Feedback



Axiom
DATA SCIENCE



Axiom
DATA SCIENCE

Data Lifecycle

DATA CREATION & QUALITY CONTROL
Scientists or Ingestion

DATA STORAGE
Workspace

DATA DESCRIPTION
Metadata Editor

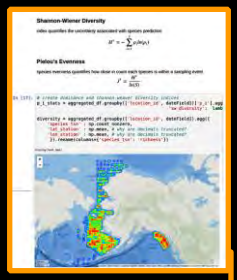
ARCHIVE & PRESERVATION
Repository submission pathway

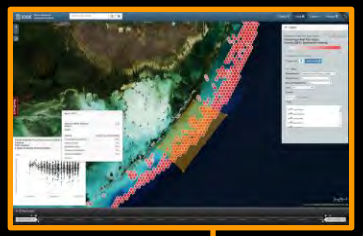
DATA ACCESS & DISCOVERY
Data portals & search catalogs

REUSE & TRANSFORMATION
Jupyter Notebook & data analyses

Axiom DATA SCIENCE

Data Lifecycle





DATA CREATION & QUALITY CONTROL
Scientists or Ingestion

DATA STORAGE
Workspace

REUSE & TRANSFORMATION
Jupyter Notebook & data analyses

Axiom DATA SCIENCE

DATA ACCESS & DISCOVERY
Data portals & search catalogs

DATA DESCRIPTION
Metadata Editor

ARCHIVE & PRESERVATION
Repository submission pathway

Data Lifecycle

DATA CREATION & QUALITY CONTROL
Scientists or Ingestion

DATA STORAGE
Workspace

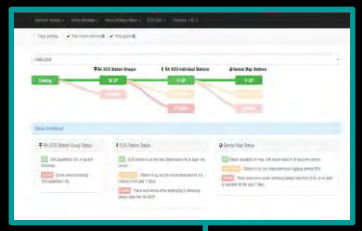
DATA DESCRIPTION
Metadata Editor

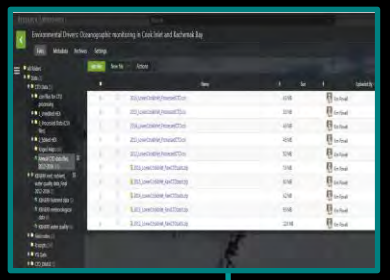
ARCHIVE & PRESERVATION
Repository submission pathway

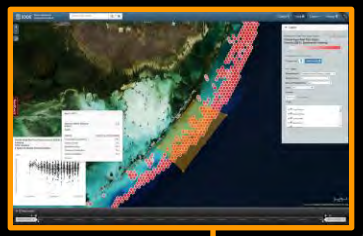
DATA ACCESS & DISCOVERY
Data portals & search catalogs

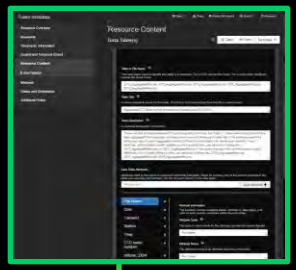
REUSE & TRANSFORMATION
Jupyter Notebook & data analyses

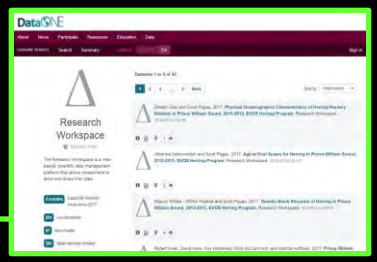
Axiom DATA SCIENCE







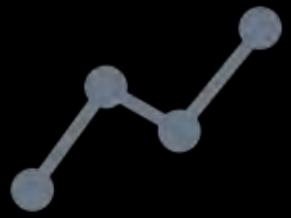




RESEARCH WORKSPACE

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

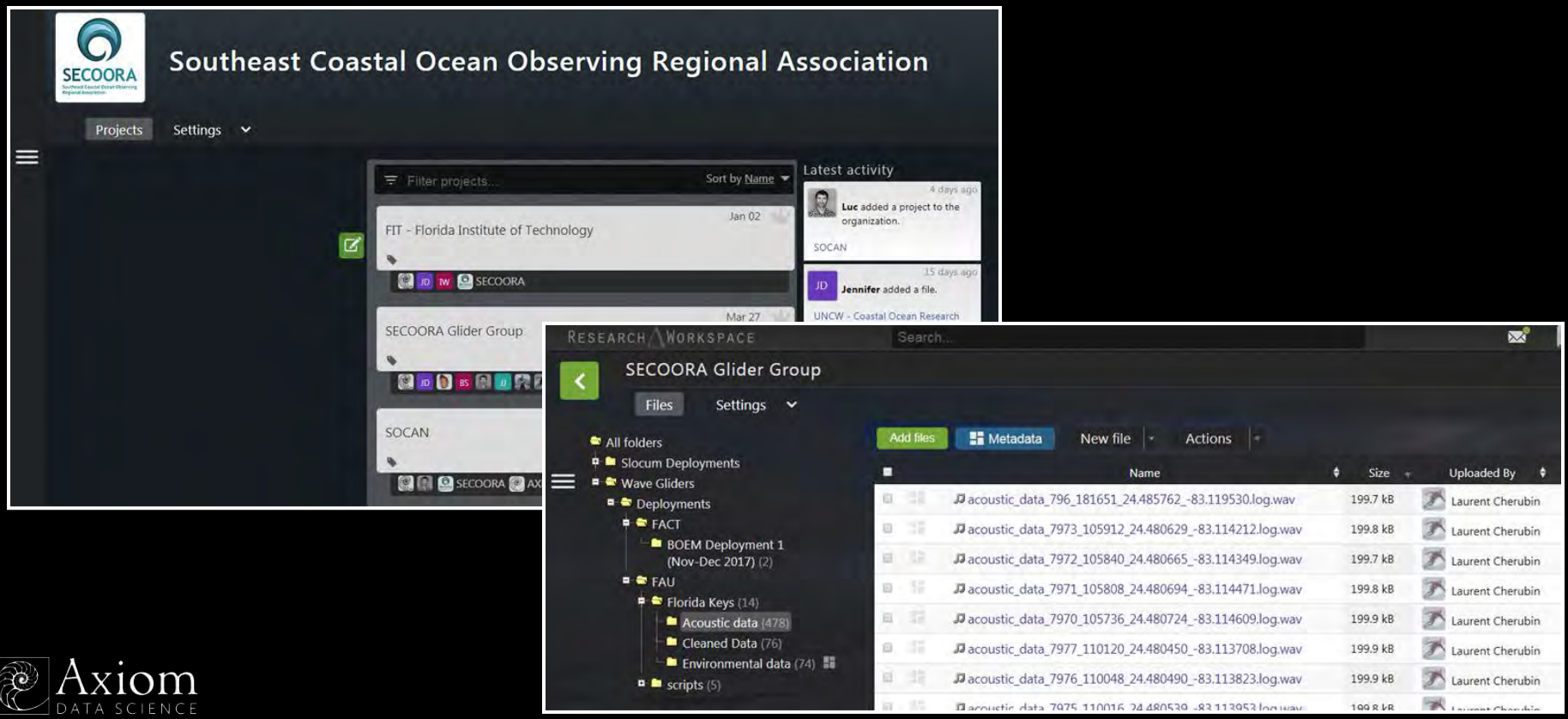
Share
Analyze
Preserve



Axiom
DATA SCIENCE

RESEARCH WORKSPACE: Store & Share

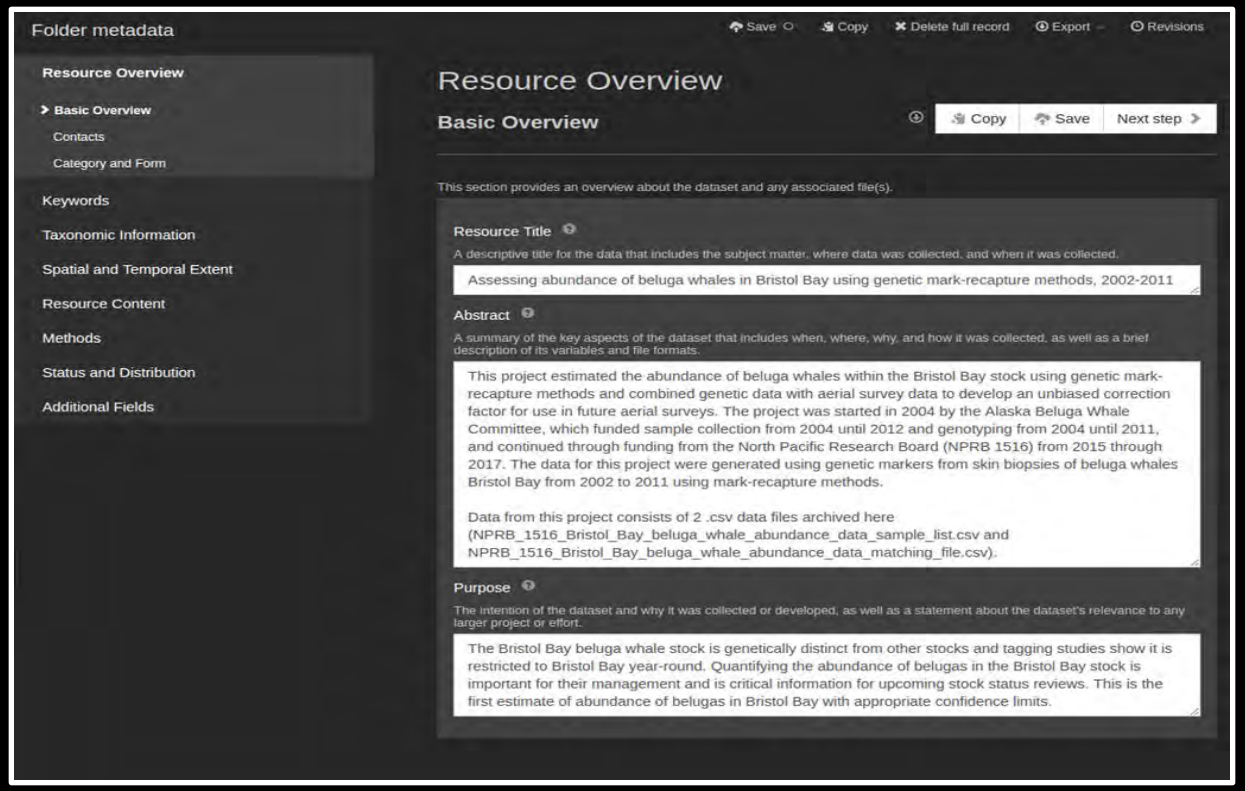
- Organize data into projects, research campaigns & organizations for coordinated data exchange
- Manage sharing through advanced security permissions



Axiom DATA SCIENCE

RESEARCH WORKSPACE : Metadata

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



Axiom DATA SCIENCE

RESEARCH WORKSPACE : Publish & Archive

Publish

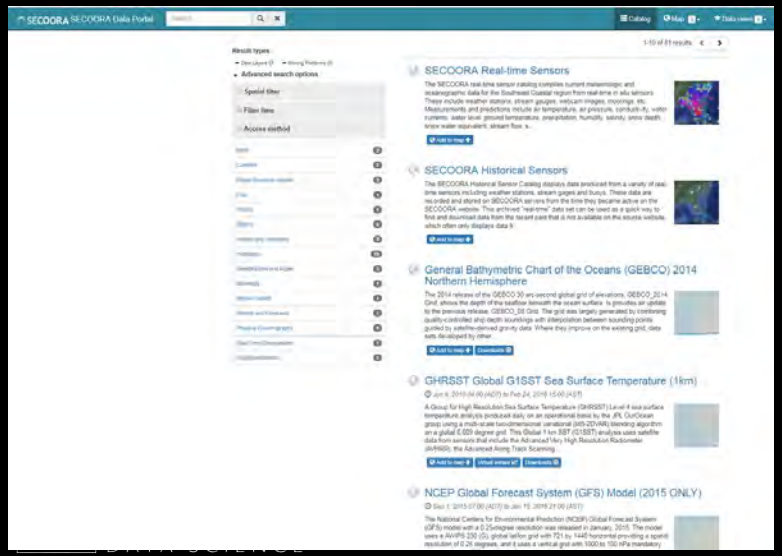
Archive

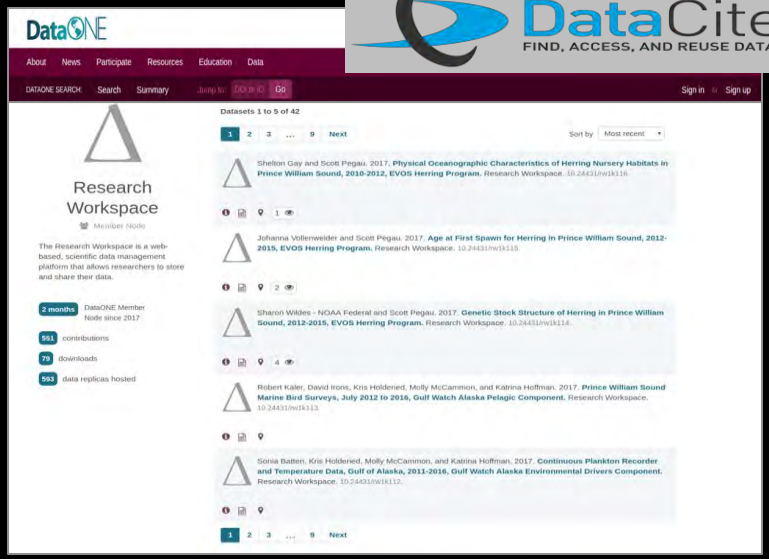
Links

SECOORA
Southeast Coastal Ocean Observing Regional Association

DataONE

DataCite
FIND, ACCESS, AND REUSE DATA

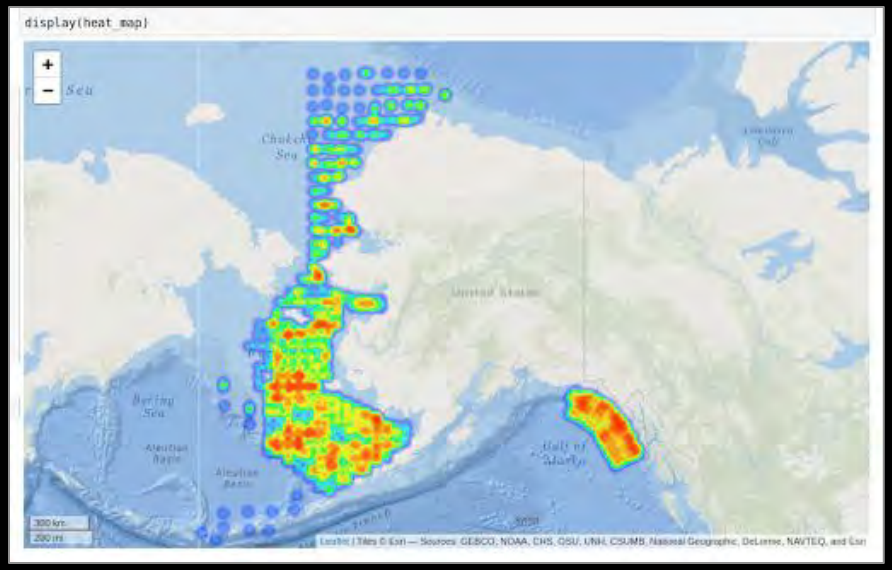




RESEARCH WORKSPACE

jupyter

- 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 toal number of individuals in the sampe

$$Dominance = max(p_i)$$

Shannon-Wiener Diversity

index quantifies the uncertainty associated with species prediction

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

Pielou's Evenness

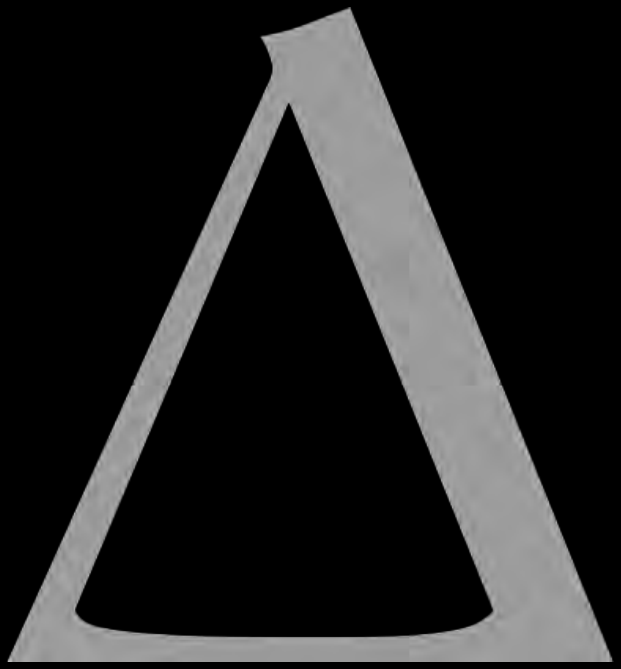
species evenness quantifies how close in count each species is within a sampling event

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

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['rich

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

Axiom DATA SCIENCE

Research Workspace & Metadata



Resources

Help Docs

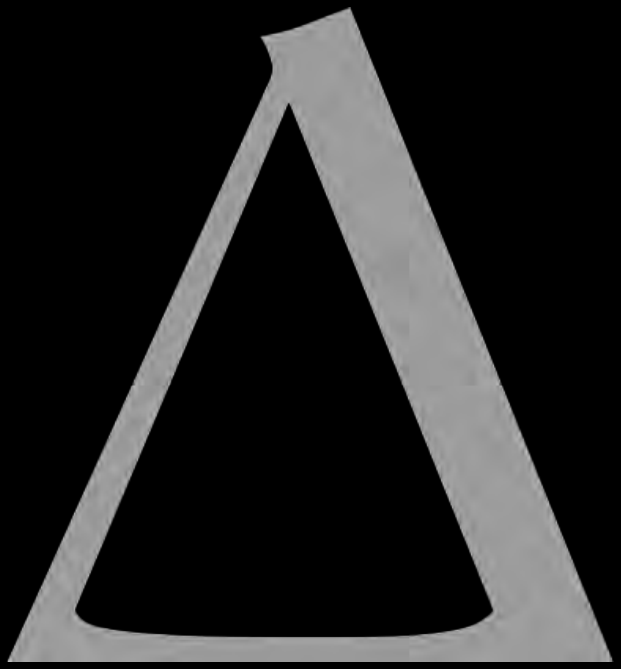
https://researchworkspace.com/help/

Technical Assistance

dmac@secoora.org

Axiom DATA SCIENCE

Research Workspace: Live Demo



https://researchworkspace.com

Axiom
DATA SCIENCE

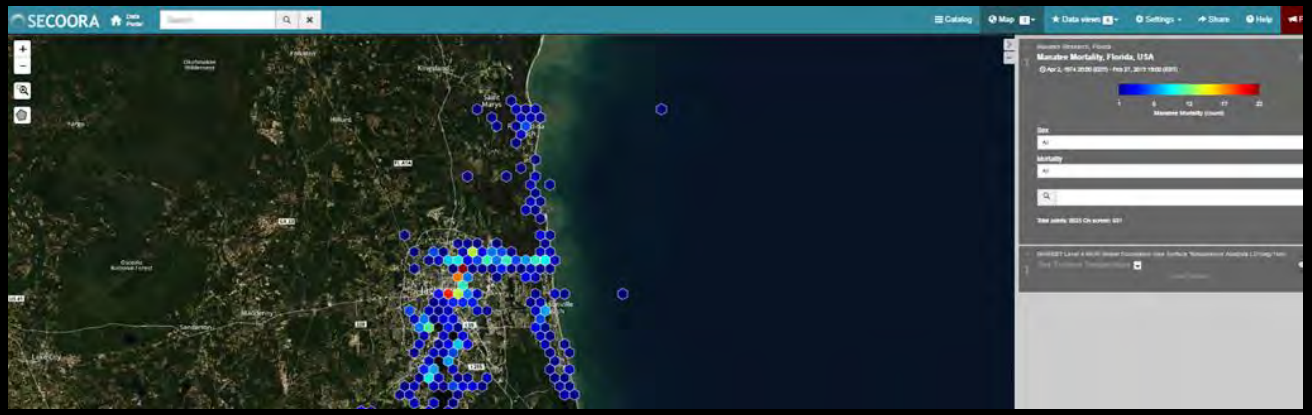
Axiom
DATA SCIENCE

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/bjiFrlI5JNy0Lwca2



Axiom DATA SCIENCE

Questions?