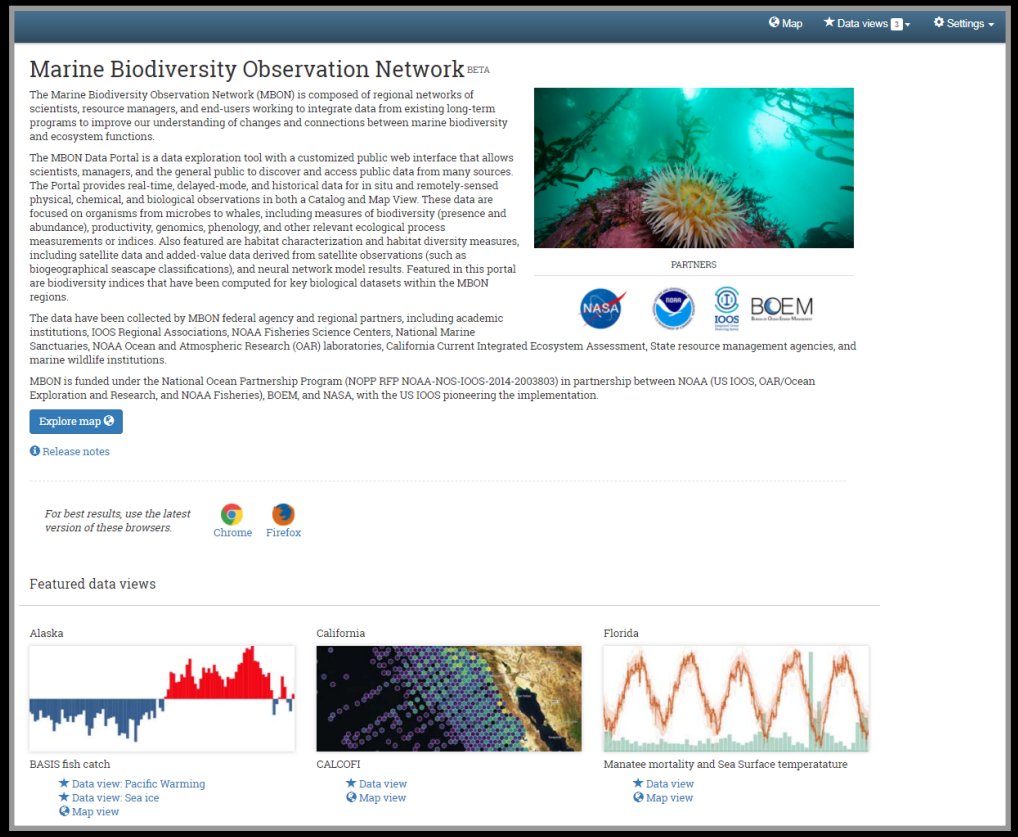
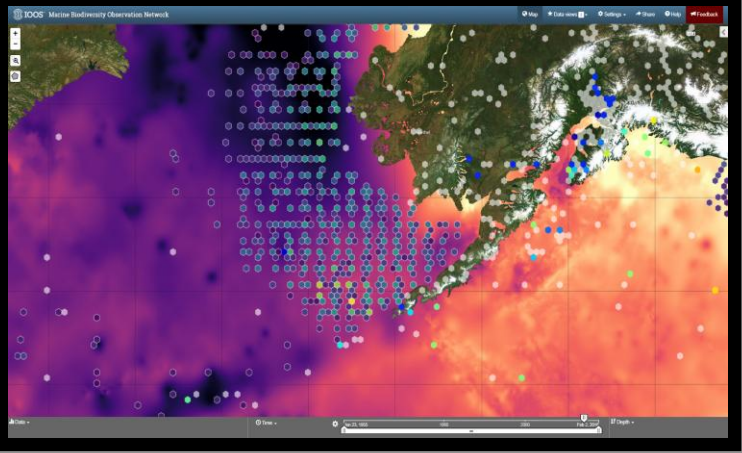
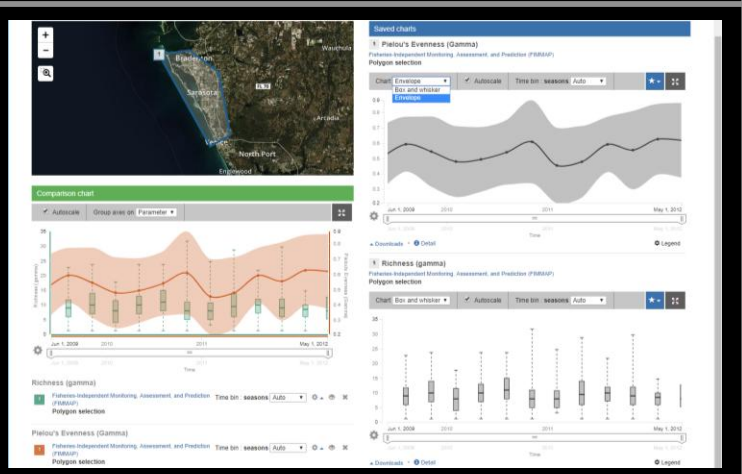
MBON Data Portal (v2.0)

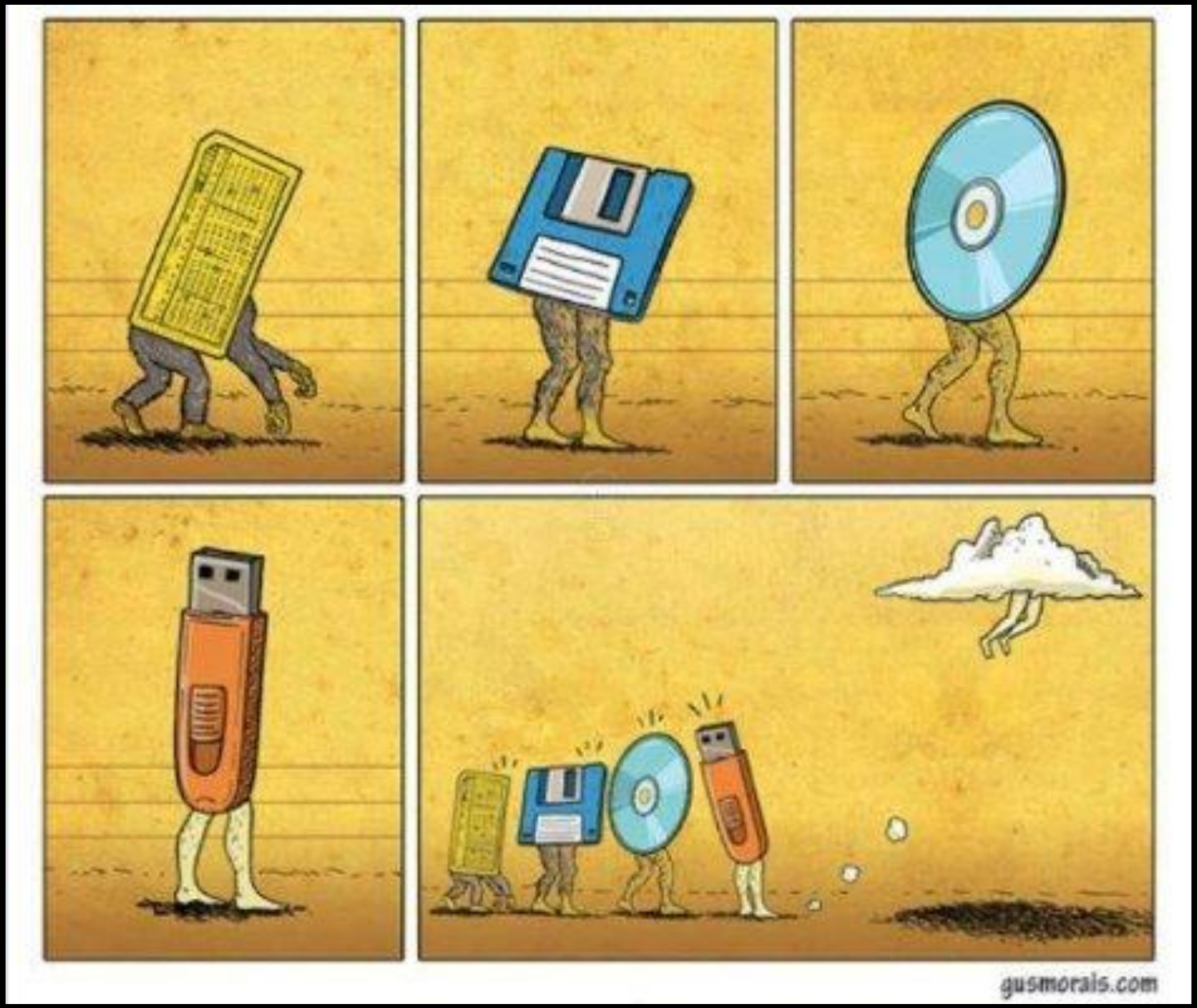






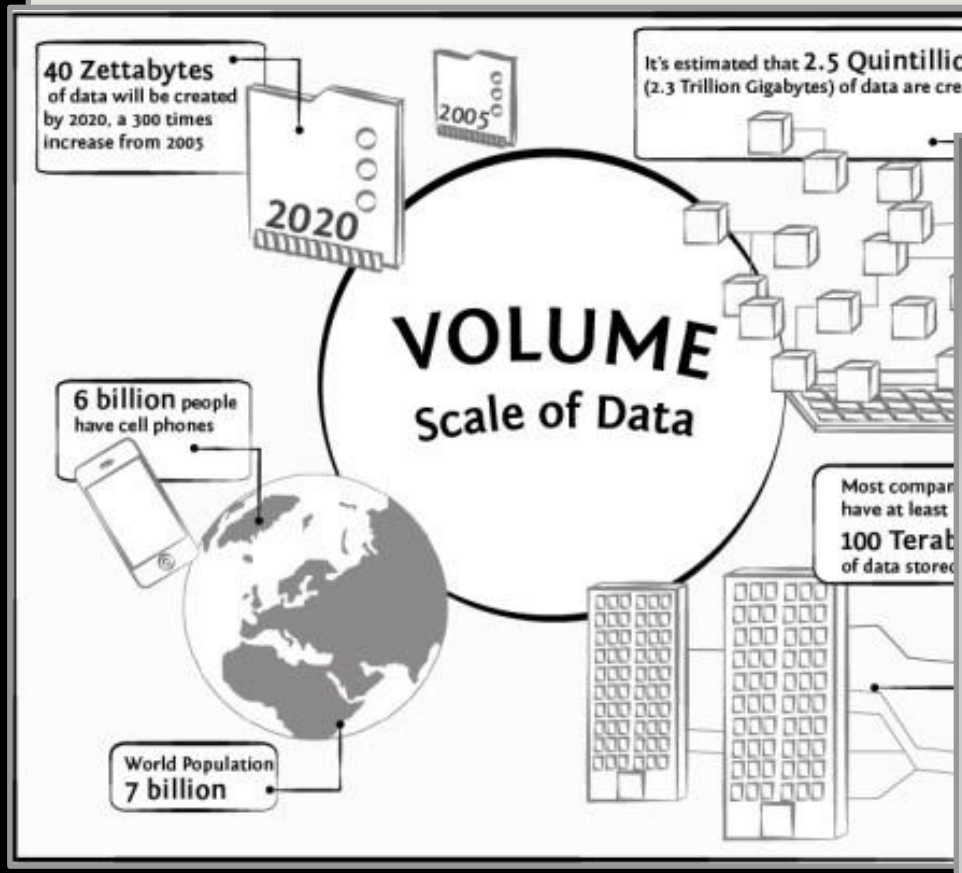
Stacey Buckelew, Axiom Data Science

Axiom Data Science | IOOS | NOAA | NASA | BOEM Bureau of Ocean Energy Management



Really "Big Data" – Data volume is increasing exponentially

By the year 2020, the digital universe will reach 44 zettabytes – that's a 10-fold increase from 2013.



Previous post
Climate change is pres
warns

NEWS BLOG

Global scientific output doubles every nine years

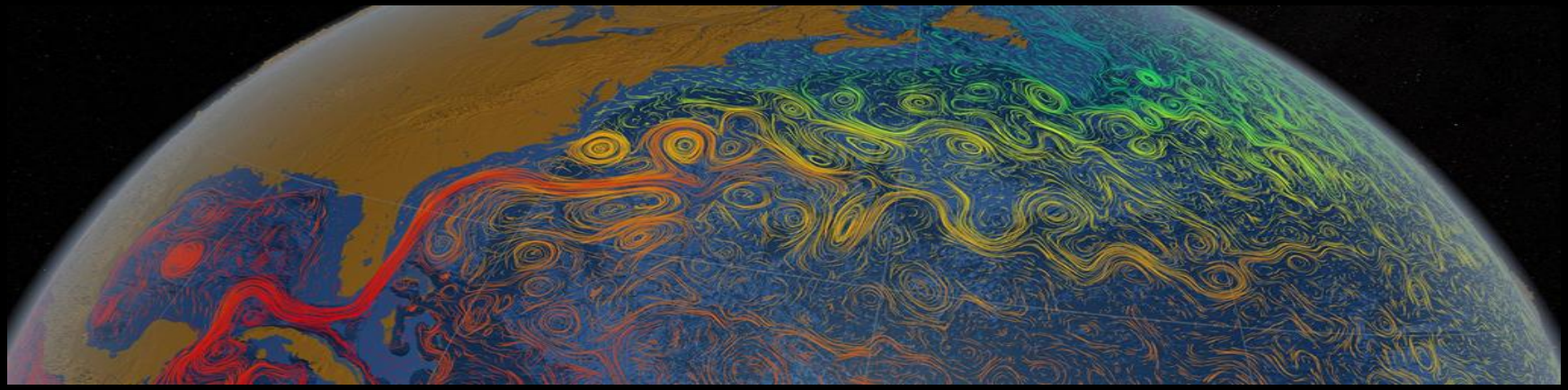
07 May 2014 | 16:46 BST | Posted by Richard Van Noorden | Category: Policy, Publishing

It's a common complaint among academics: today's researchers are publishing too much, too fast. B
how fast is the mass of scientific output actually growing?

Axiom DATA SCIENCE

Data Management Changes

- Data explosion is reforming data management and governance
- Leveraging existing data systems in different ways- *architecture, storage, databases*
- Emphasis on open data as crucial to reliable science
- Integration, visualization, and dissemination are tools, not by-products



Axiom
DATA SCIENCE

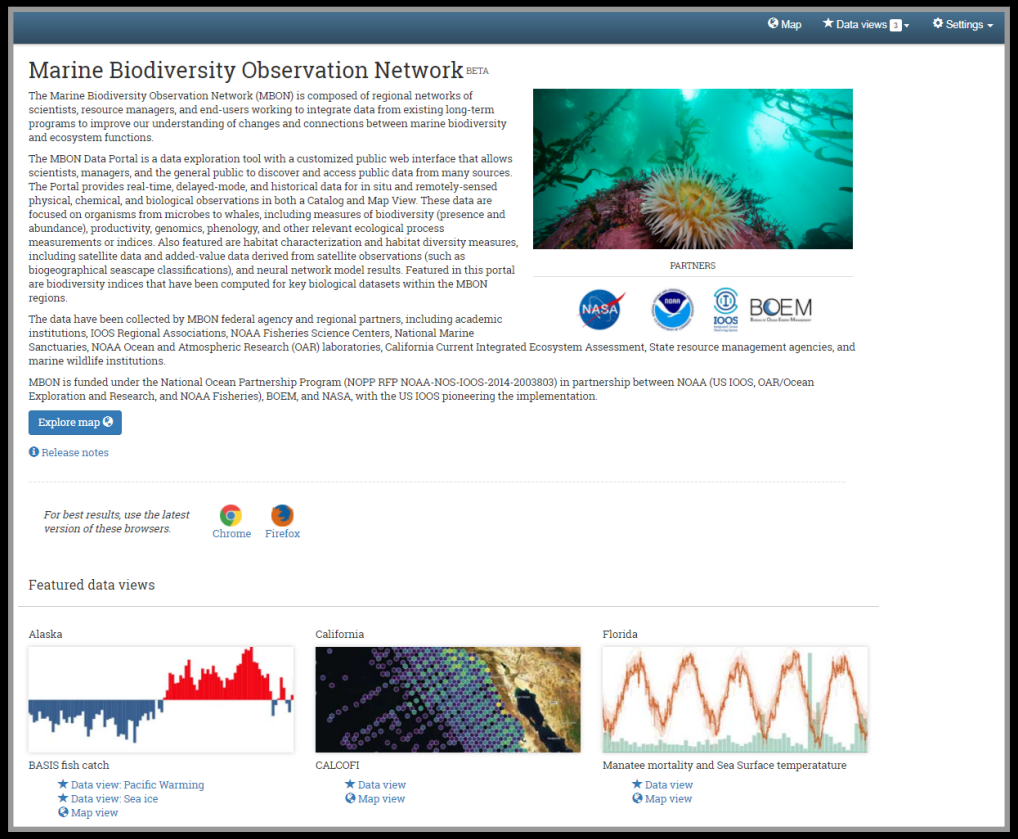
OCEAN DATA IS BIG DATA

- Shared cyberinfrastructure approach
  - Leverage applications, systems & hardware across partners
- Community developed software, standards and protocols
  - Standardize systems and interfaces across partners
- Scalable compute and storage infrastructure (HPC)
  - 5 petabytes storage; 3,000 processing cores

Axiom DATA SCIENCE

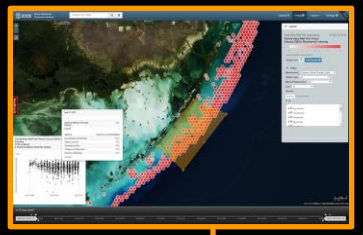
MBON Data Portal (v2.0)

https://mbon.ioos.us



Customizable data environment allowing scientists, managers, and the public others to discover, interact, and access biologically-relevant data from different sources

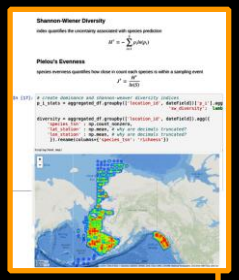
Axiom DATA SCIENCE

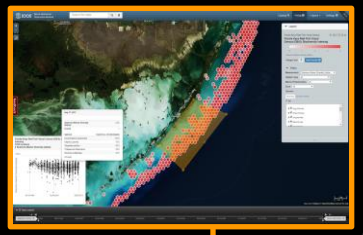


REUSE & TRANSFORMATION

Jupyter Notebook & data analyses

Axiom DATA SCIENCE

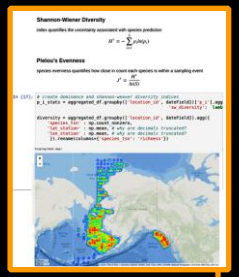


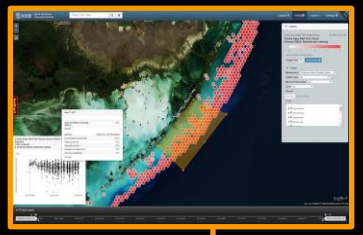


REUSE & TRANSFORMATION

Jupyter Notebook & data analyses

Axiom DATA SCIENCE



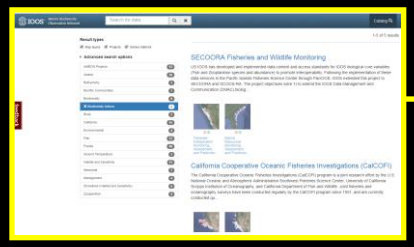


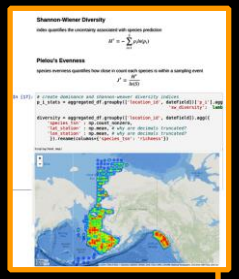
REUSE & TRANSFORMATION

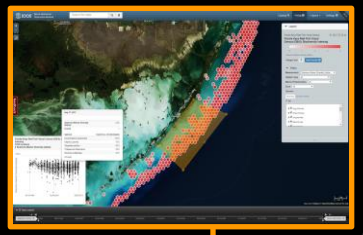
Jupyter Notebook & data analyses

DATA ACCESS & DISCOVERY

Data portals & search catalogs

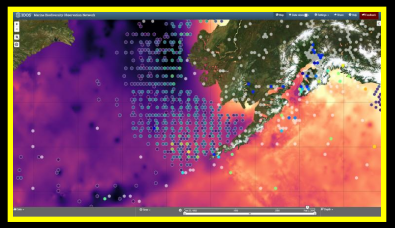






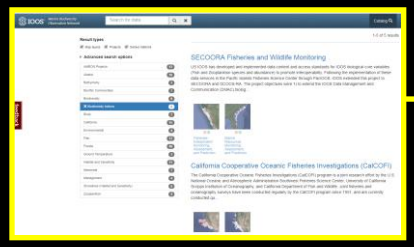
REUSE & TRANSFORMATION

Jupyter Notebook & data analyses

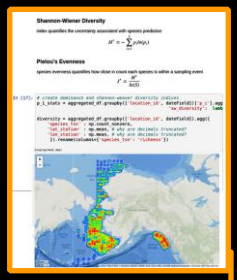


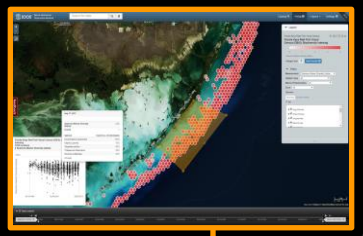
DATA ACCESS & DISCOVERY

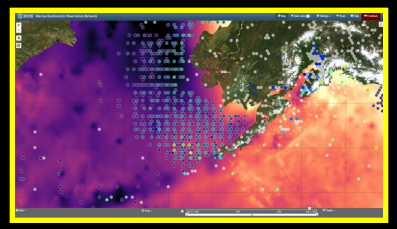
Data portals & search catalogs

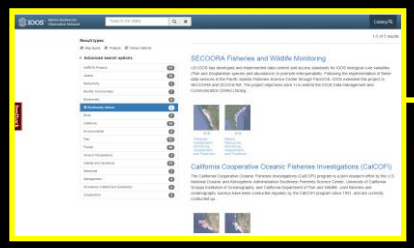


Data Lifecycle









DATA CREATION & QUALITY CONTROL
Scientists or Ingestion

DATA STORAGE
Workspace

DATA DESCRIPTION
Metadata Editor

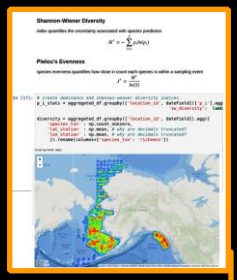
ARCHIVE & PRESERVATION
Repository submission pathway

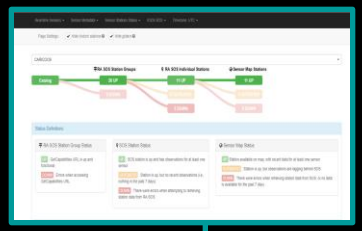
DATA ACCESS & DISCOVERY
Data portals & search catalogs

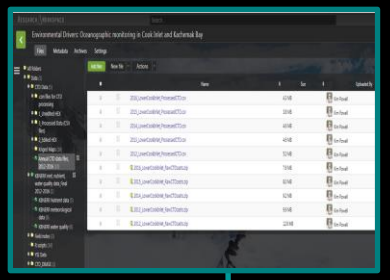
REUSE & TRANSFORMATION
Jupyter Notebook & data analyses

Axiom DATA SCIENCE

Data Lifecycle

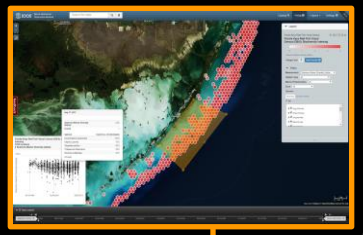






DATA CREATION & QUALITY CONTROL

Scientists or Ingestion

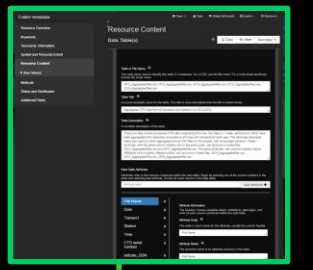


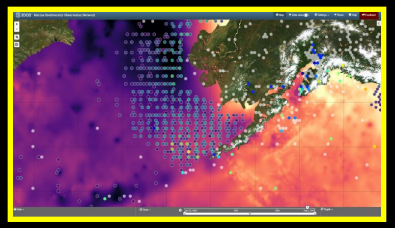
REUSE & TRANSFORMATION

Jupyter Notebook & data analyses

DATA STORAGE

Workspace



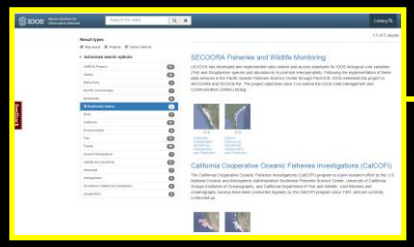


DATA ACCESS & DISCOVERY

Data portals & search catalogs

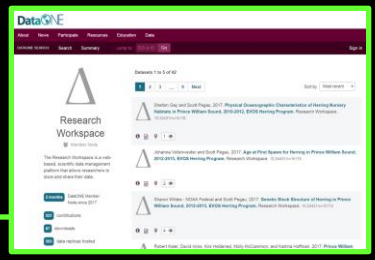
DATA DESCRIPTION

Metadata Editor



ARCHIVE & PRESERVATION

Repository submission pathway

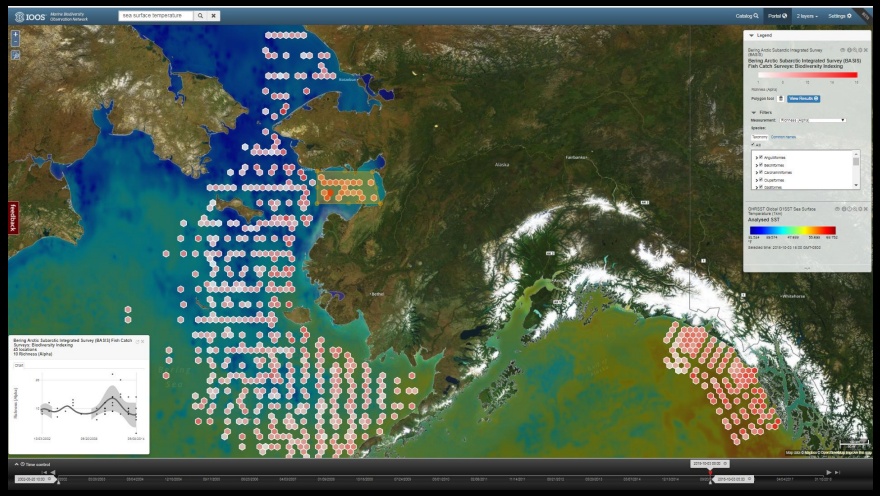


Axiom DATA SCIENCE

IOOS

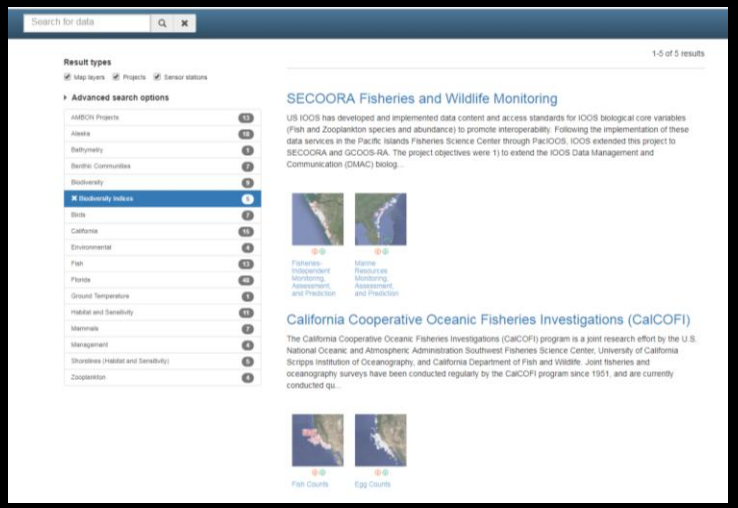
Map

Integrate & visualize data from many sources



Catalog

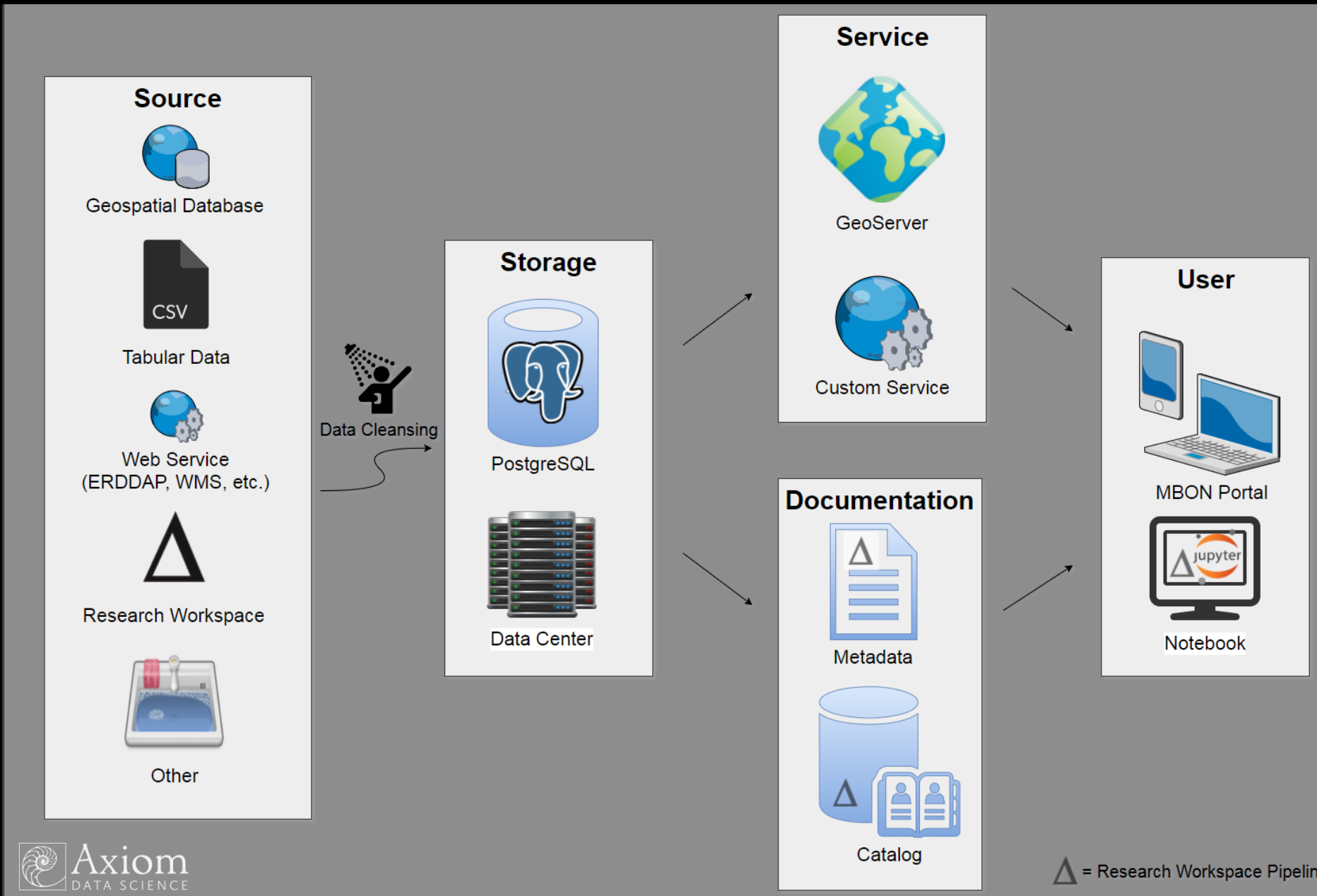
Search , metadata, & data download



Data Views

Rapidly assimilate & compare different data streams

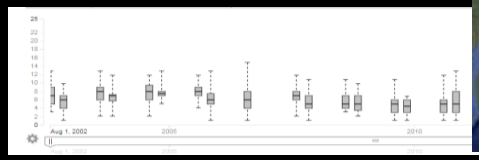
MBON Data Pipeline



Data Types

Biodiversity

count, richness, diversity indices

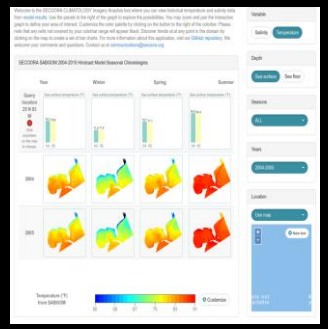


Platforms

moorings, shore stations

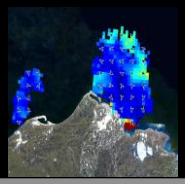
Products

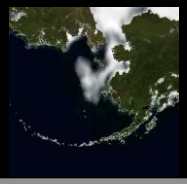
skill assessment, shoreline change, etc.

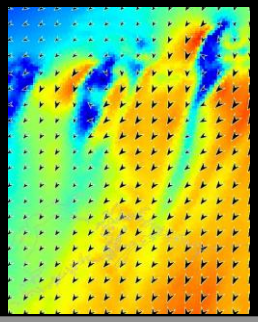


Grids

models, satellite, radar

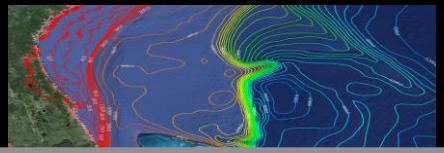




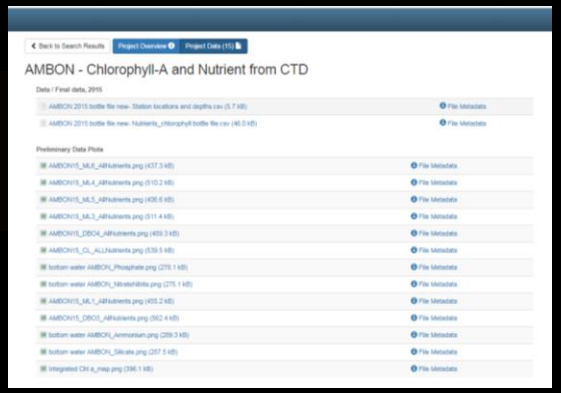


GIS

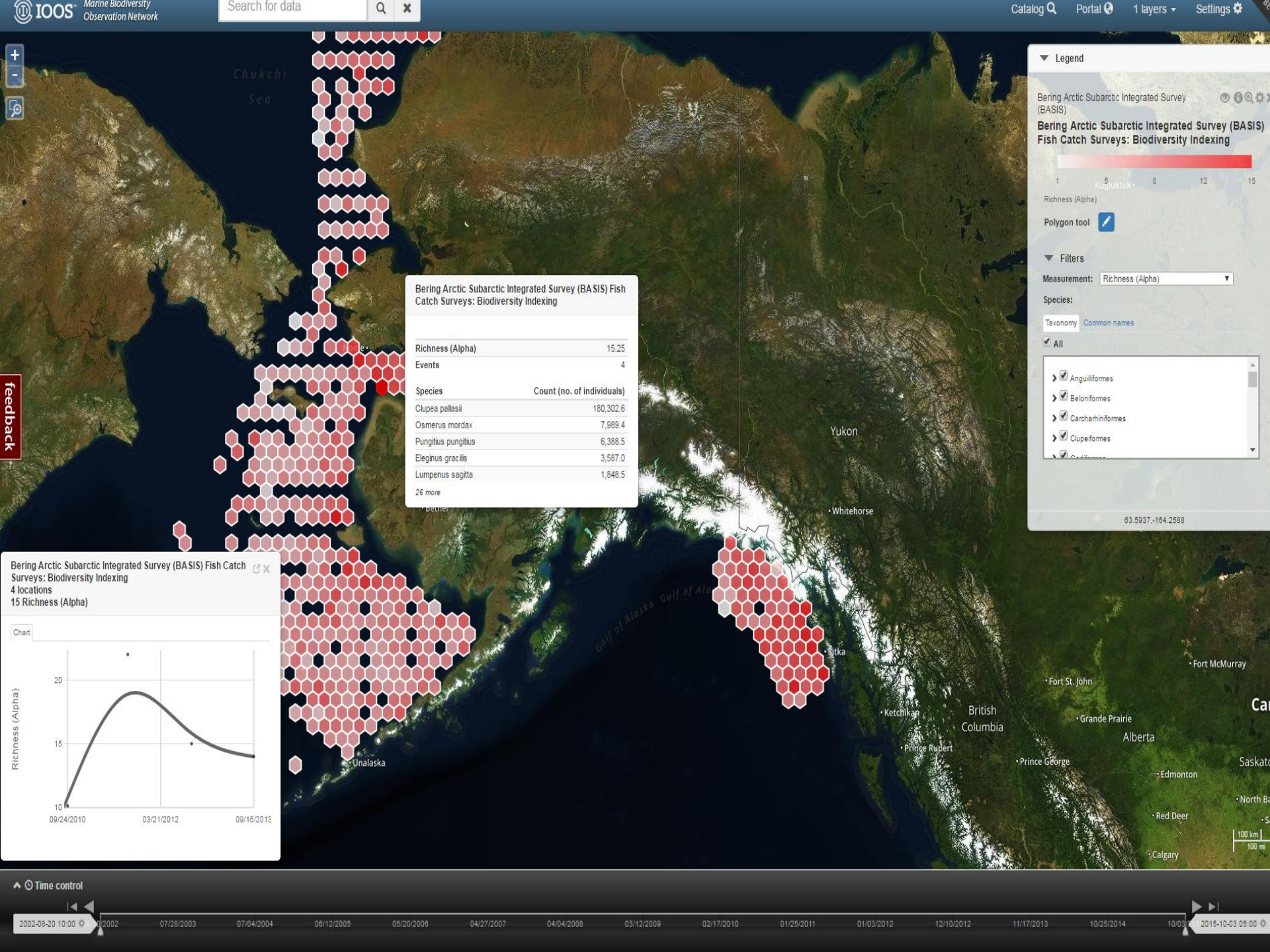
Habitat types, bathymetry, fishing zones, etc.

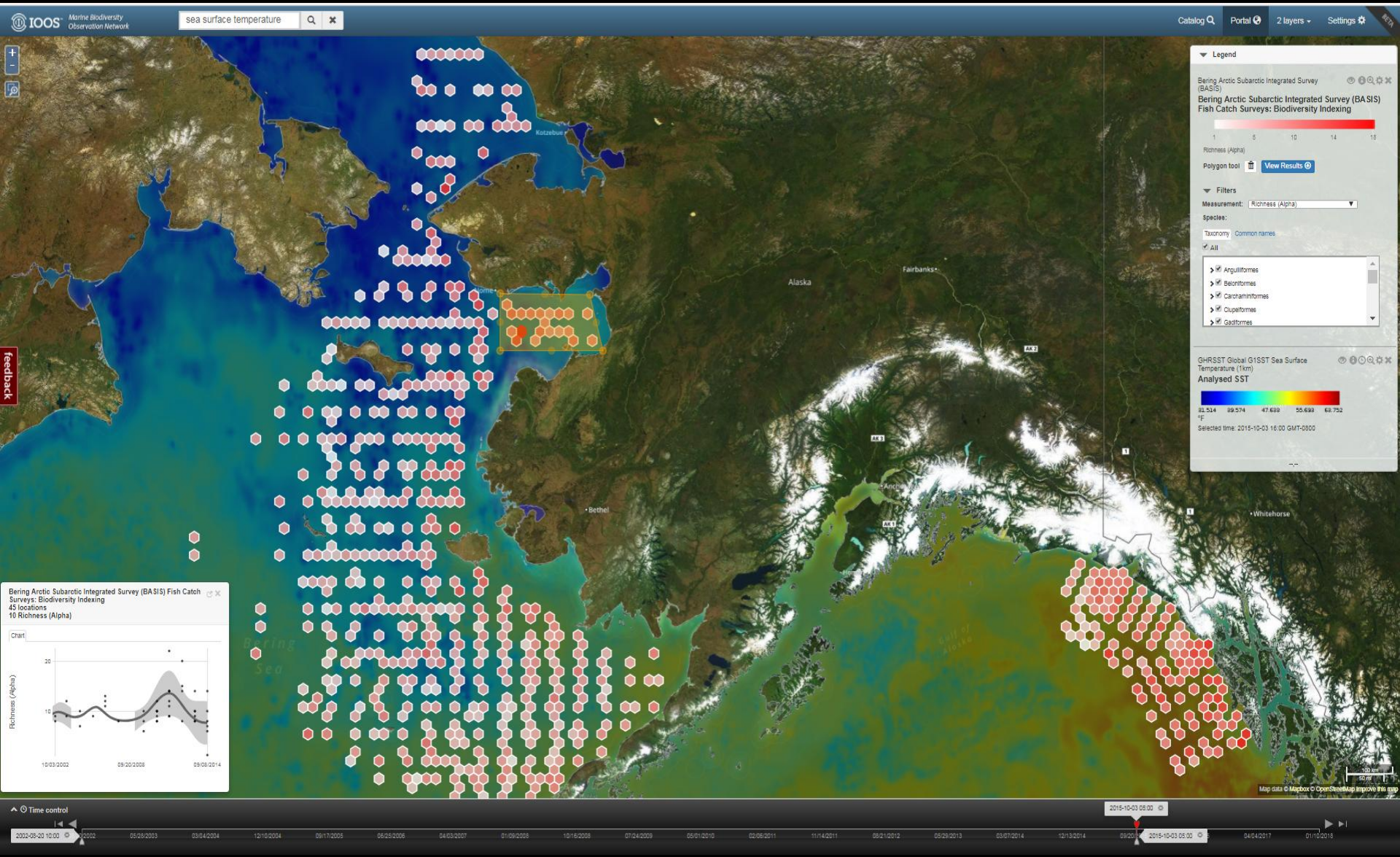


RESEARCH WORKSPACE



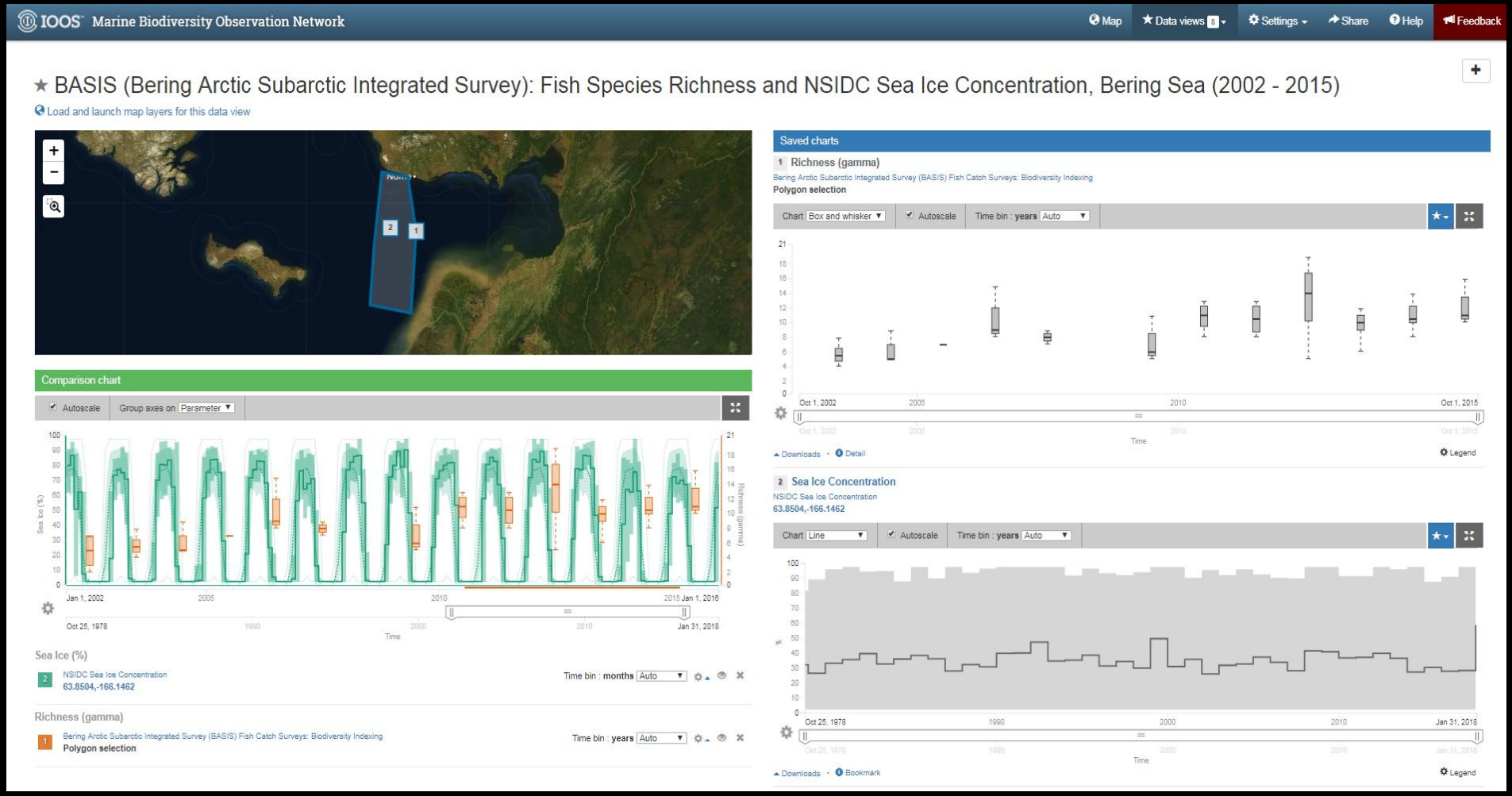
Axiom DATA SCIENCE


IOOS Marine Biodiversity Observation Network
Search for data
Catalog
Portal
1 layers
Settings
Chukchi Sea
Yukon
Whitehorse
Gulf of Alaska
Unalaska
Sitka
Ketchikan
Prince Rupert
British Columbia
Prince George
Fort St. John
Fort McMurray
Grande Prairie
Alberta
Edmonton
Red Deer
Calgary
Kuguktuk
Legend
Bering Arctic Subarctic Integrated Survey (BASIS)
Bering Arctic Subarctic Integrated Survey (BASIS) Fish Catch Surveys: Biodiversity Indexing
1
5
8
12
15
Richness (Alpha)
Polygon tool
Filters
Measurement: Richness (Alpha)
Species:
Taxonomy
Common names
All
Anguilliformes
Beloniformes
Carcharhiniformes
Clupeiformes
63.5937,-164.2588
Bering Arctic Subarctic Integrated Survey (BASIS) Fish Catch Surveys: Biodiversity Indexing
Richness (Alpha) 15.25
Events 4
Species
Count (no. of individuals)
Clupea pallasii 180,302.6
Osmerus mordax 7,989.4
Pungitius pungitius 6,388.5
Eleginus gracilis 3,587.0
Lumpenus sagitta 1,848.5
26 more
Bering Arctic Subarctic Integrated Survey (BASIS) Fish Catch Surveys: Biodiversity Indexing
4 locations
15 Richness (Alpha)
Chart
Richness (Alpha)
20
15
10
09/24/2010
03/21/2012
09/16/2013
feedback
Time control
2002-08-20 10:00
07/28/2003
07/04/2004
06/12/2005
05/20/2006
04/27/2007
04/04/2008
03/12/2009
02/17/2010
01/25/2011
01/03/2012
12/10/2012
11/17/2013
10/25/2014
2015-10-03 05:00
100 km
100 mi



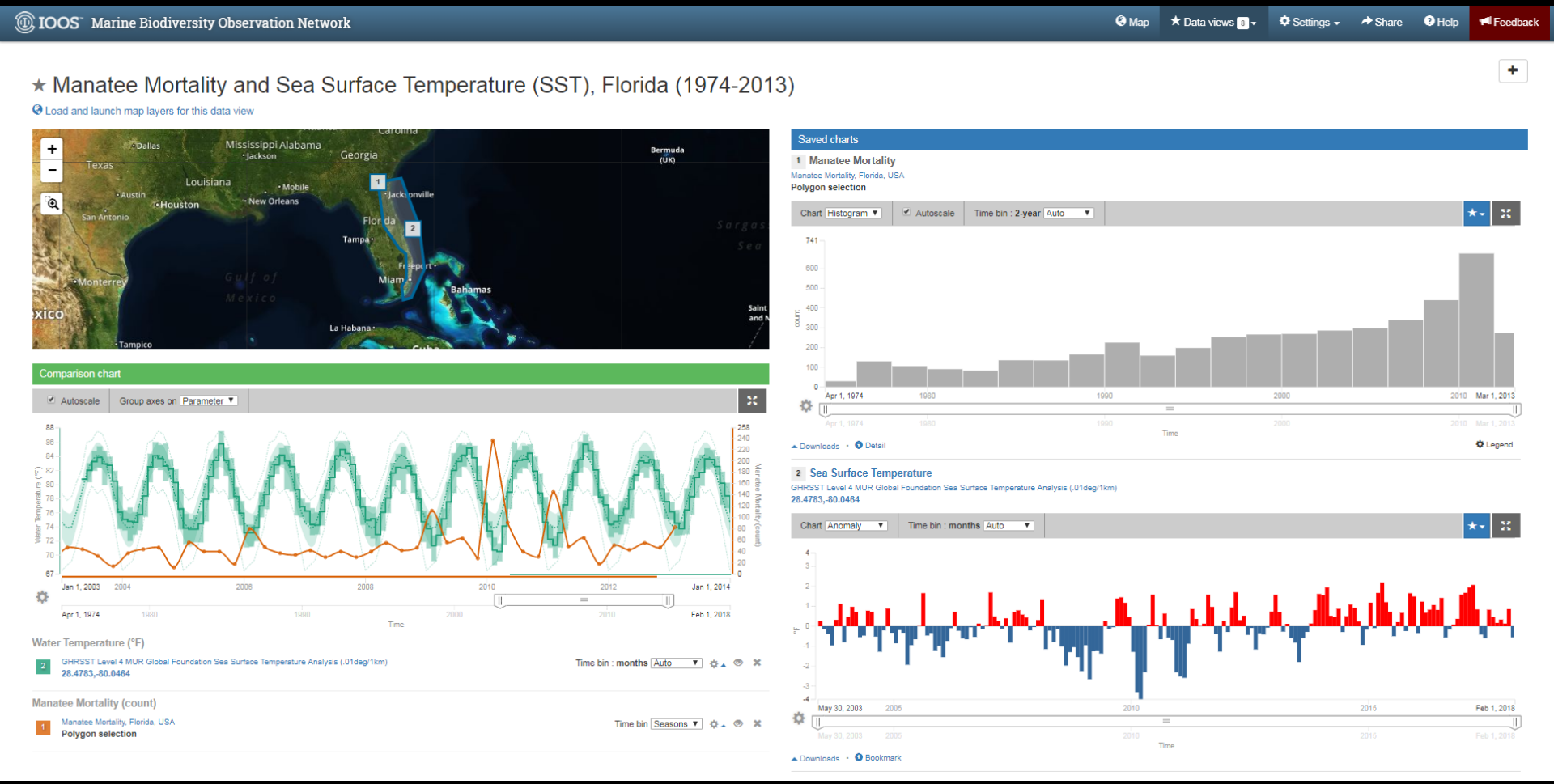
Axiom DATA SCIENCE

Custom Data Views (‘dashboards’)



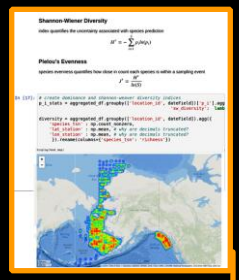
Axiom DATA SCIENCE

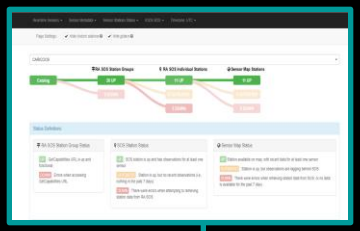
Custom Data Views (‘dashboards’)

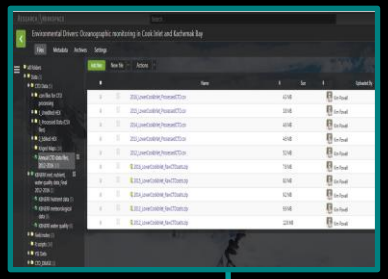


Axiom DATA SCIENCE

Data Lifecycle

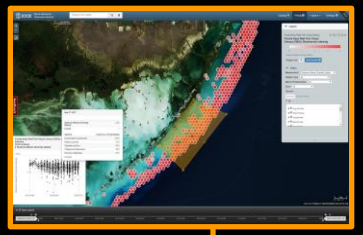






DATA CREATION & QUALITY CONTROL

Scientists or Ingestion

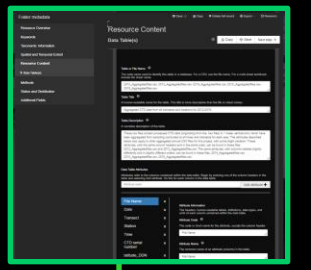


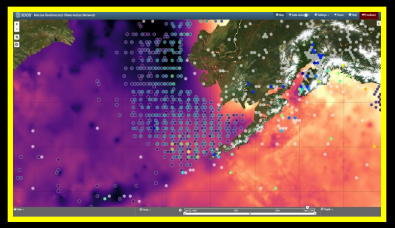
REUSE & TRANSFORMATION

Jupyter Notebook & data analyses

DATA STORAGE

Workspace



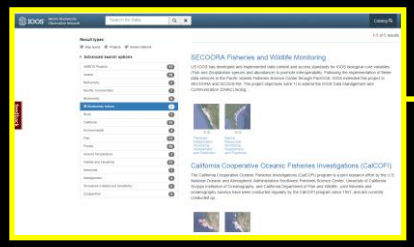


DATA ACCESS & DISCOVERY

Data portals & search catalogs

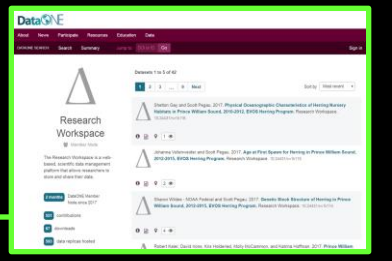
DATA DESCRIPTION

Metadata Editor



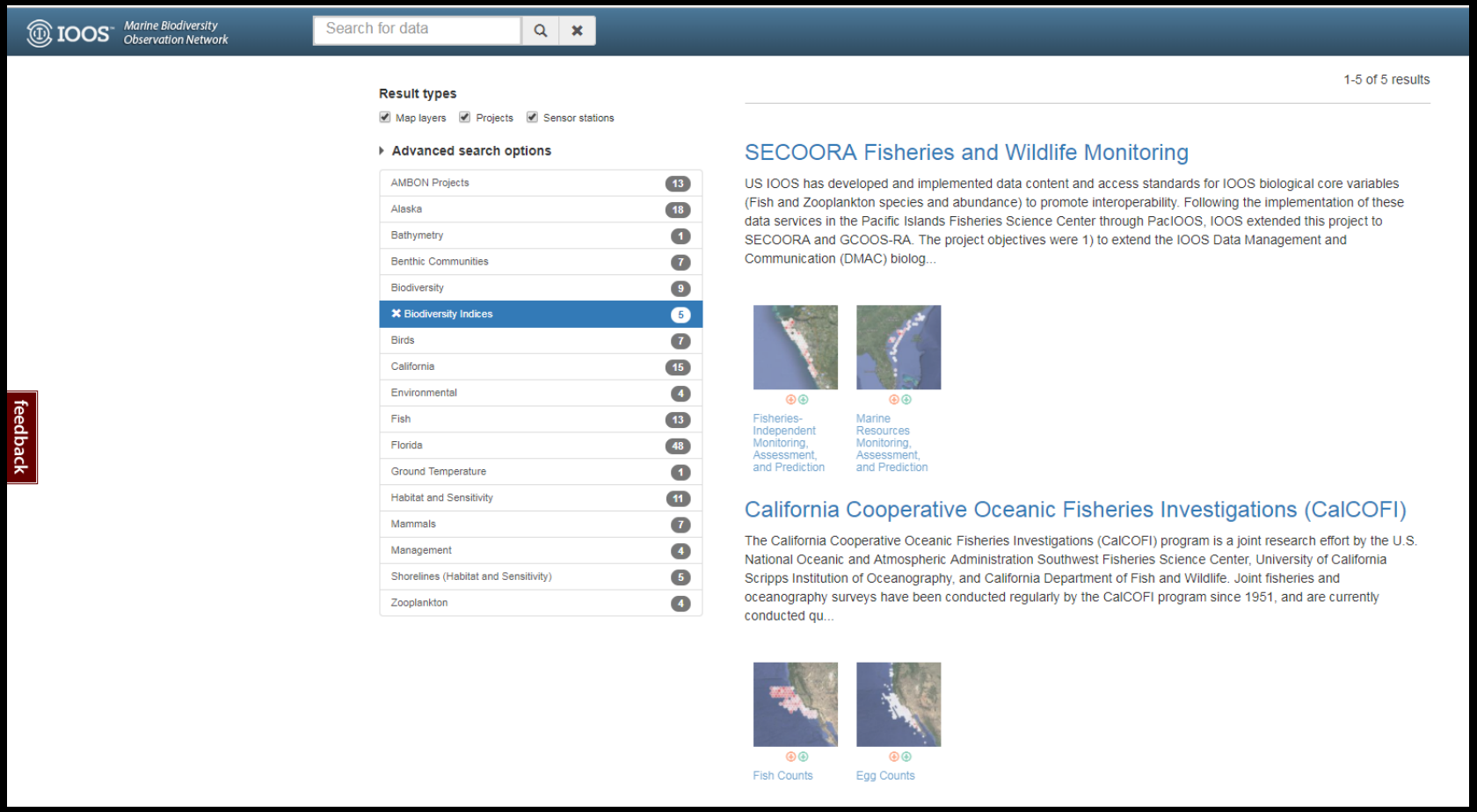
ARCHIVE & PRESERVATION

Repository submission pathway



Axiom DATA SCIENCE

MBON Data Catalog



Axiom DATA SCIENCE

[MBON Portal (http://mbon.ioos.us)](http://mbon.ioos.us)

Data Catalog

Content Pages

- Summary
- Data availability
- License / usage
- Variable / layer preview
- Contact information
- Original-source link

Back to Search Results

Rockfish Recruitment and Ecosystem Assessment Survey

The Fisheries Ecology Division (FED), of the Southwest Fisheries Science Center (SWFSC) has conducted a midwater trawl survey off central California since 1983 with the primary goal of developing pre-recruit indices for young-of-the-year (YOY) rockfish (Sebastes spp.). The survey also samples numerous other components of the epipelagic micronekton, including other YOY groundfish (such as Pacific hake, Merluccius productus, and sanddab, Citharichthys spp), coastal pelagic fishes (such as Pacific sardine, Sardinops sagax, and northern anchovy, Engraulis mordax) and other forage species.

Rockfish Recruitment and Ecosystem Assessment Survey

Download ▾ | Portal | +

- **Metadata URL:** http://oceanview.pfeg.noaa.gov/erddap/in...
- **Date Range:** 05/13/1990 11:52 - 06/12/2015 19:27

The Fisheries Ecology Division (FED), of the Southwest Fisheries Science Center (SWFSC) has conducted a midwater trawl survey off central California since 1983 with the primary goal of developing pre-recruit indices for young-of-the-year (YOY) rockfish (Sebastes spp.). The survey also samples numerous other components of the epipelagic micronekton, including other YOY groundfish (such as Pacific hake, Merluccius productus, and sanddab, Citharichthys spp), coastal pelagic fishes (such as Pacific sardine, Sardinops sagax, and northern anchovy, Engraulis mordax) and other forage species.

Additional details regarding the survey methods and results are described in Ralston et al. (2015) and Sakuma et al. Ralston, S., J.C. Field and K.S. Sakuma. 2015. Longterm variation in a central California pelagic forage assemblage. Journal of Marine Systems 146: 26-37, http://dx.doi.org/10.1016/j.jmarsys.2014.06.013. Sakuma, K.M., J.C. Field, B.B. Marinovic, C.N. Carrion, N.J. Mantua and S. Ralston. In revision. Anomalous epipelagic micronekton assemblage patterns in the neritic waters of the California Current in spring 2015 during a period of extreme ocean conditions. CalCOFI Reports.

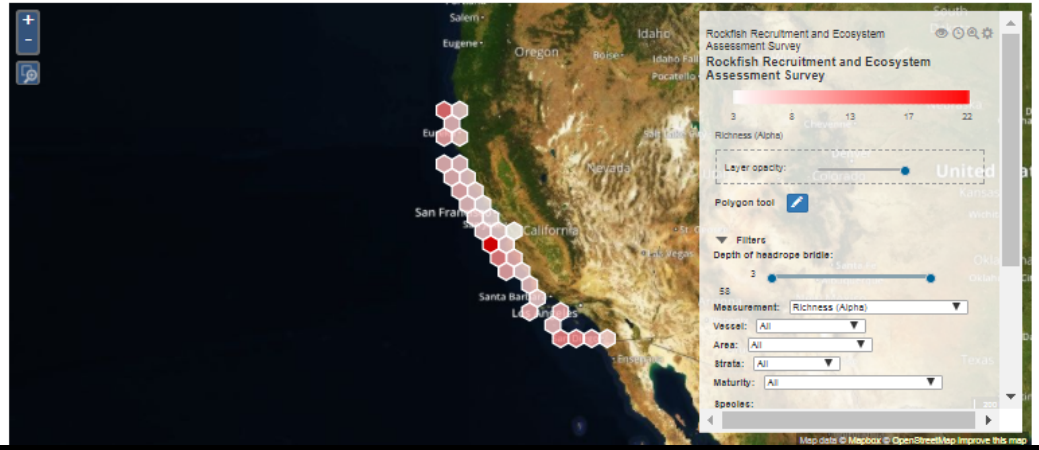
Filter options: *Biodiversity indexes are not normalized!*

Alpha: Average of event values within the selected area
Gamma: Group all events within the selected area and treat them as a single sample
Beta: Gamma/Alpha ratio

Richness: Count of distinct species
% Dominance (Berger-Parker): Numerical importance of the most abundant species
Shannon-Wiener Diversity: This index quantifies the uncertainty associated with species prediction
Pielou's Evenness: Species evenness quantifies how close in count each species is within a sampling event

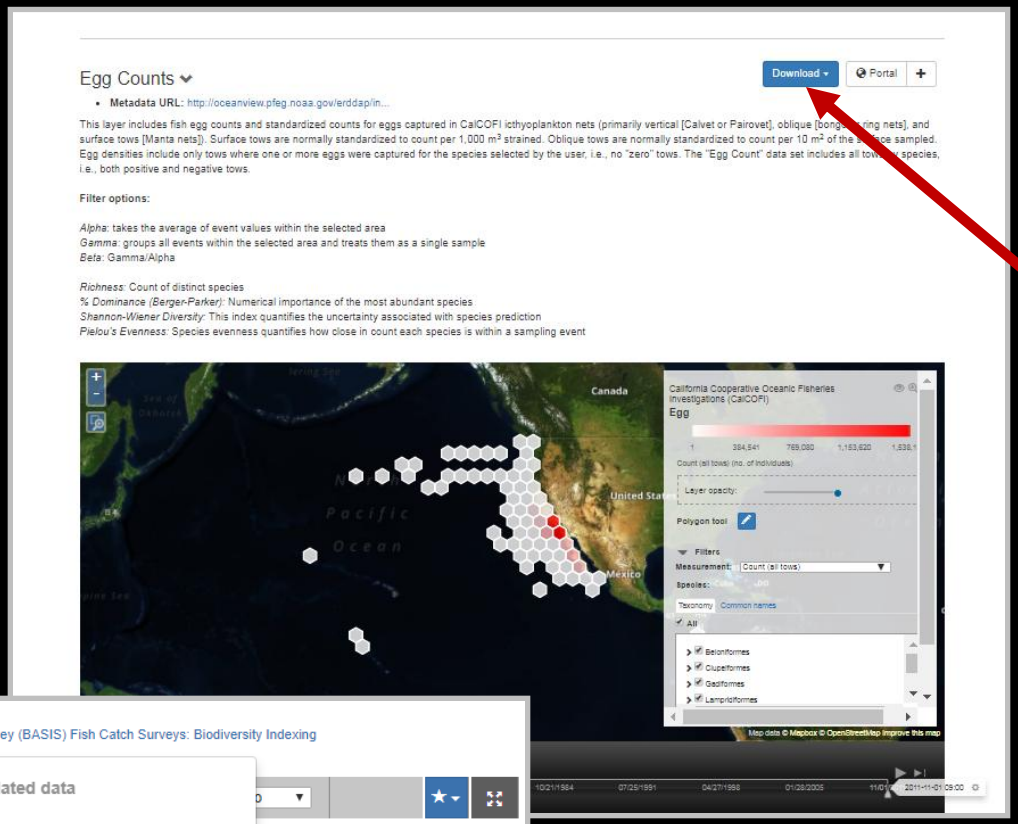
Usage notes

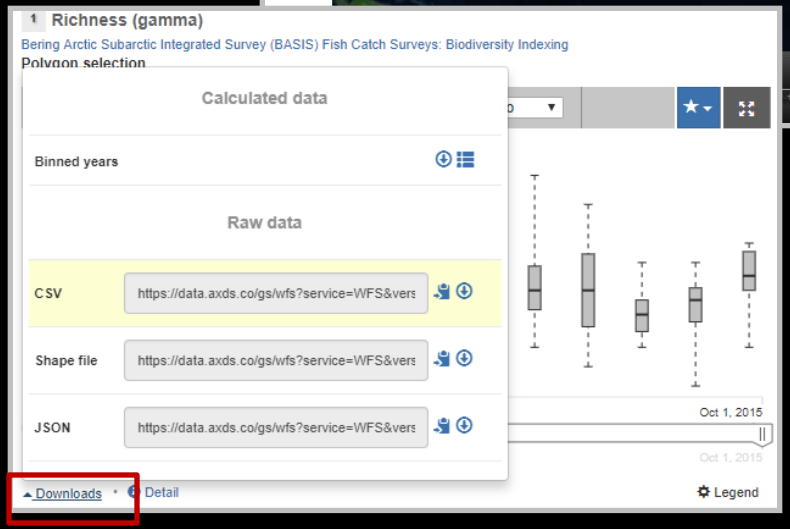
The data may be used and redistributed for free but is not intended for legal use, since it may contain inaccuracies. Neither the data Contributor, ERD, NOAA, nor the United States Government, nor any of their employees or contractors, makes any warranty, express or implied, including warranties of merchantability and fitness for a particular purpose, or assumes any legal liability for the accuracy, completeness, or usefulness, of this information.



Axiom DATA SCIENCE

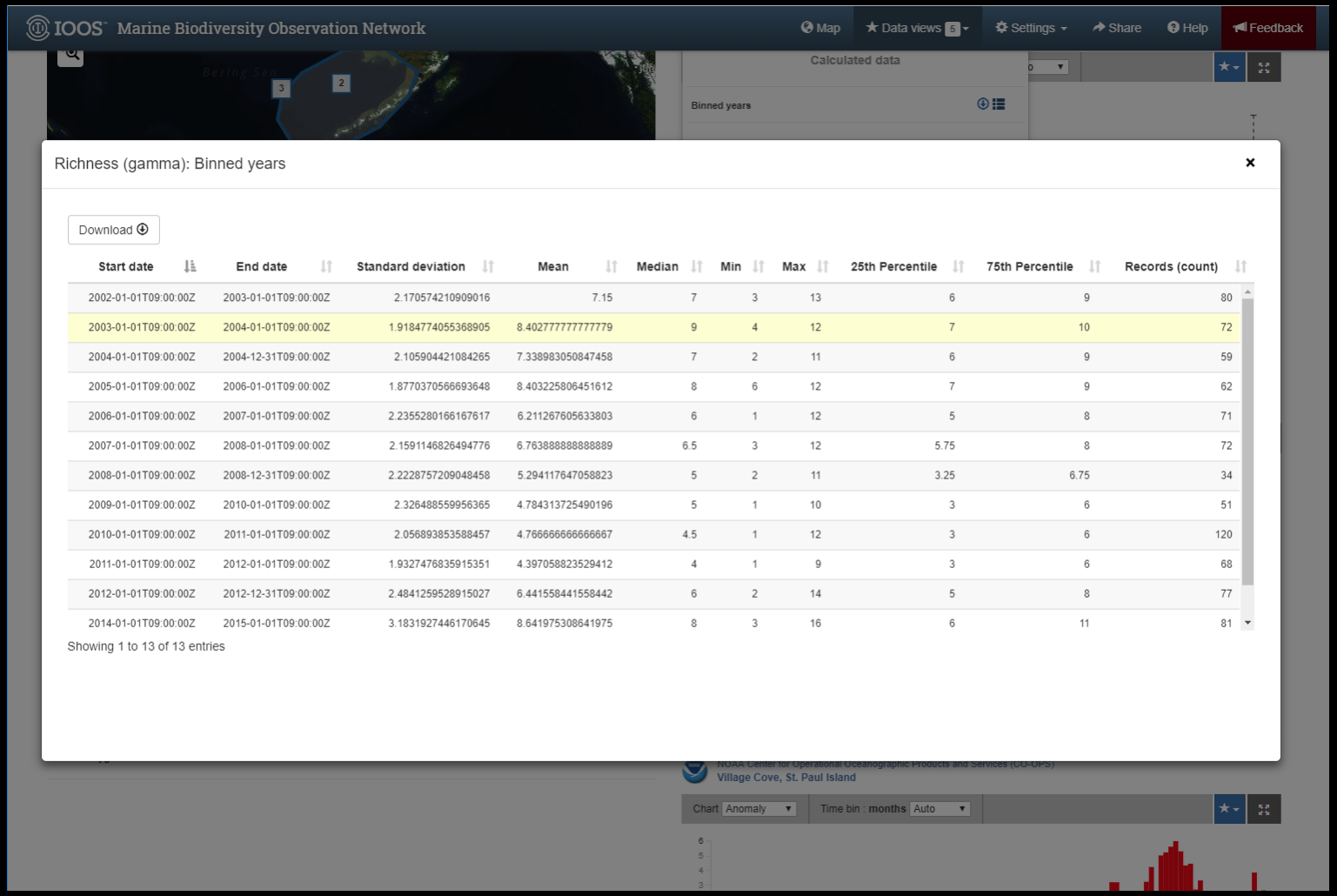
Downloads Using Interoperability





- ncWMS
- Shapefile
- CSV
- THREDDS
- netCDF
- OPenDAP
- ERDDAP

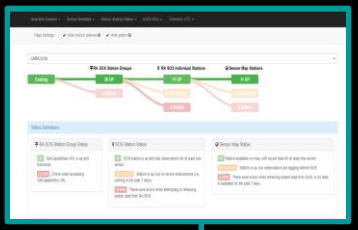
Downloads Using Interoperability

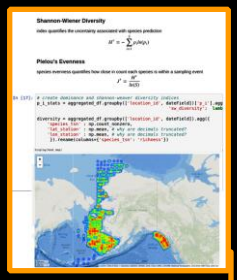


Axiom DATA SCIENCE

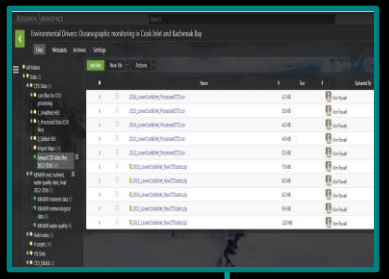
Data Lifecycle

DATA CREATION & QUALITY CONTROL
Scientists or Ingestion

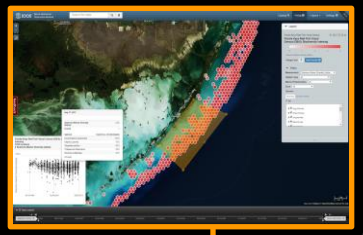




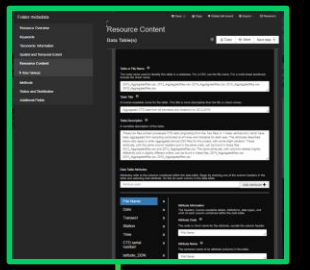
DATA STORAGE
Workspace



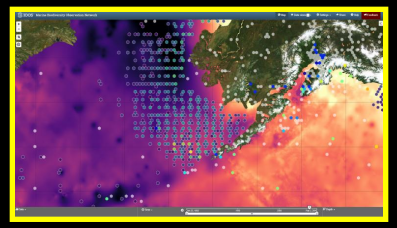
REUSE & TRANSFORMATION
Jupyter Notebook & data analyses



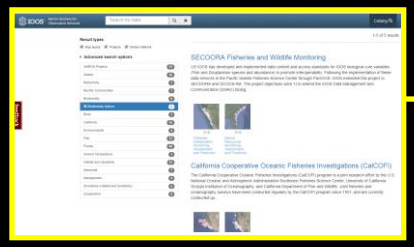
DATA DESCRIPTION
Metadata Editor

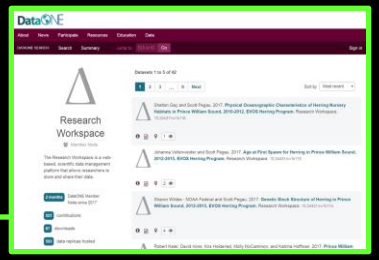


DATA ACCESS & DISCOVERY
Data portals & search catalogs



ARCHIVE & PRESERVATION
Repository submission pathway





Axiom DATA SCIENCE

RESEARCH WORKSPACE

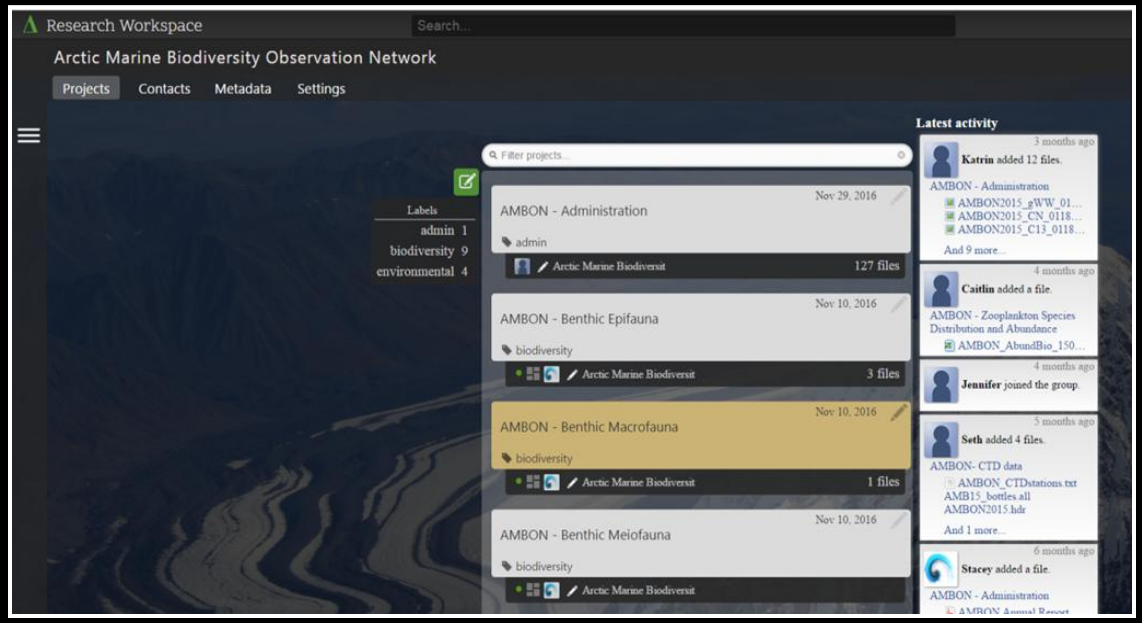
~web-based platform for collaboratively managing science projects through the entire data lifecycle~

Share
Analyze
Preserve

Axiom
DATA SCIENCE

RESEARCH WORKSPACE

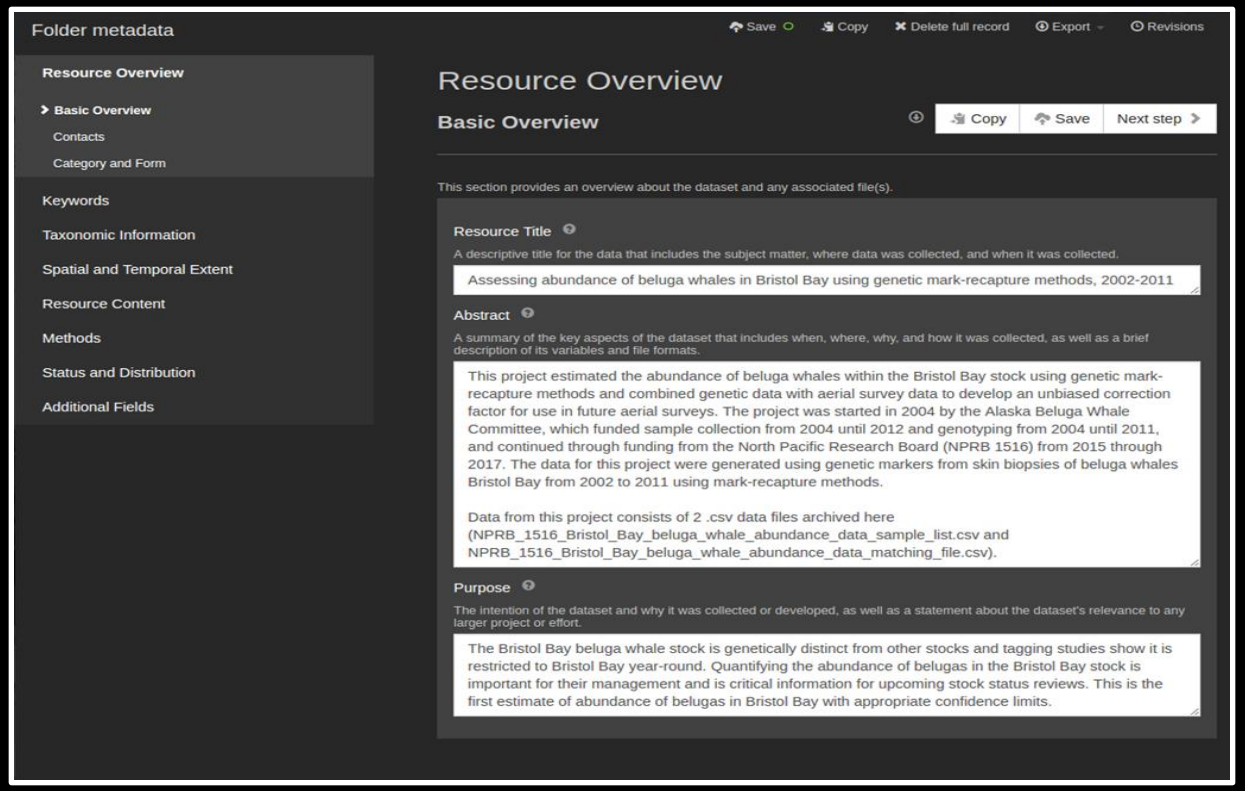
- Organize into projects, research campaigns and organizations
- Coordinate data exchange across networks, groups, programs
- ISO 19110/19115-2 standards metadata editor
- Execute server side R and Python numeric workflows (Jupyter) on uploaded data AND any data in Axiom CI stack
- Archive pathway to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



Axiom DATA SCIENCE

RESEARCH WORKSPACE : Metadata

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

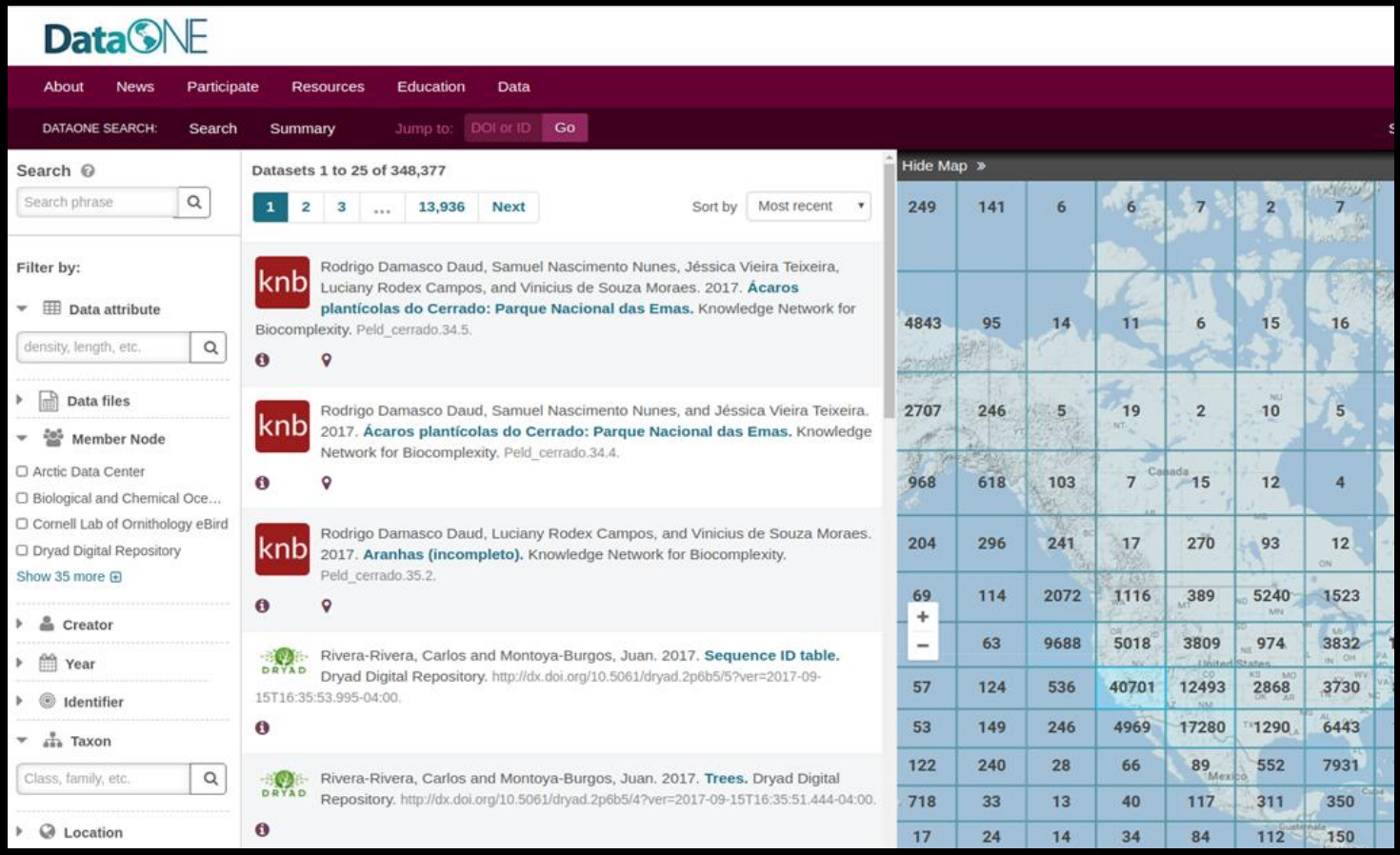


Axiom DATA SCIENCE

RESEARCH WORKSPACE

DataONE

Dataset Review → Flag for Archive → Reserve DOI → Update Metadata → Package Dataset → Push to RW MN → CN Index and Sync

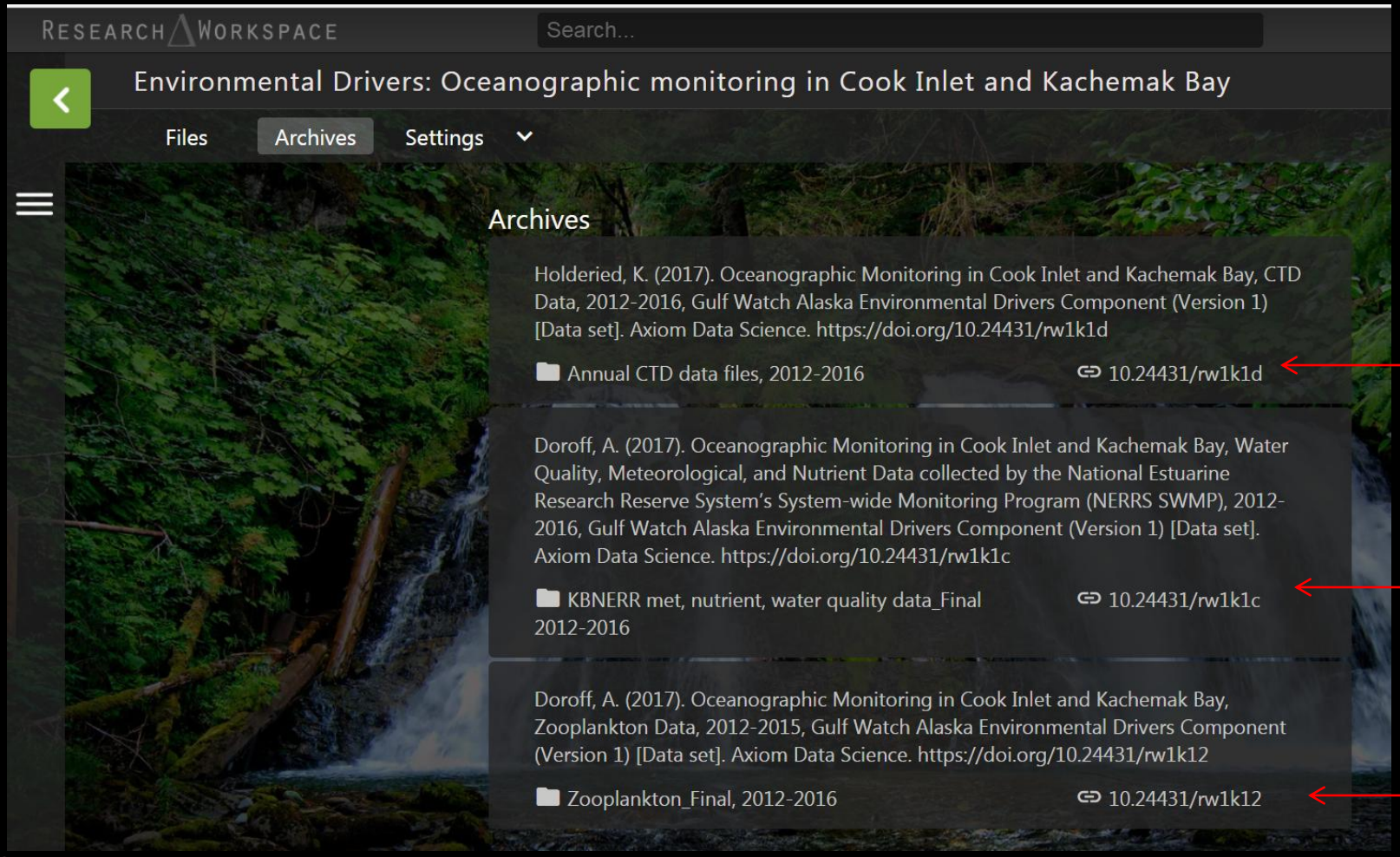


Axiom DATA SCIENCE

RESEARCH WORKSPACE

DataCite
FIND, ACCESS, AND REUSE DATA

- Locate, identify and cite research data

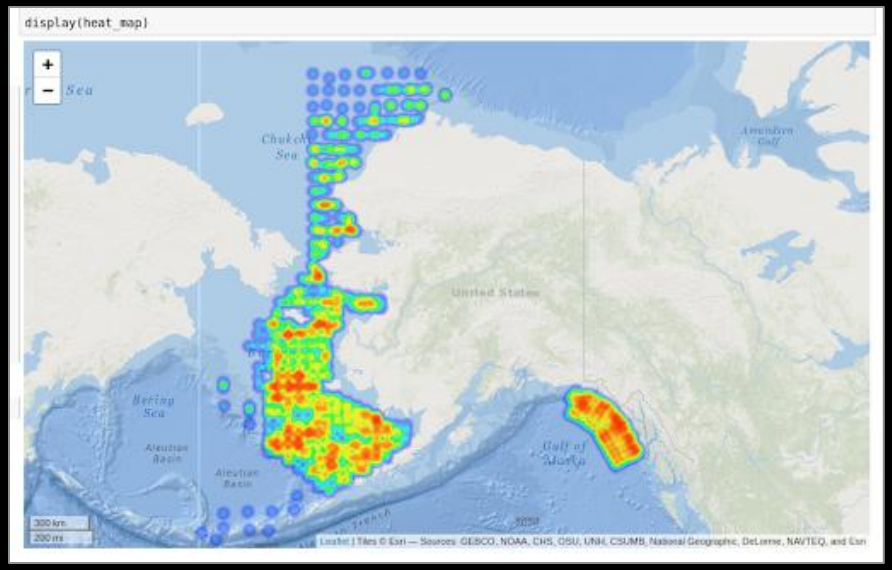


Axiom
DATA SCIENCE

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



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

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

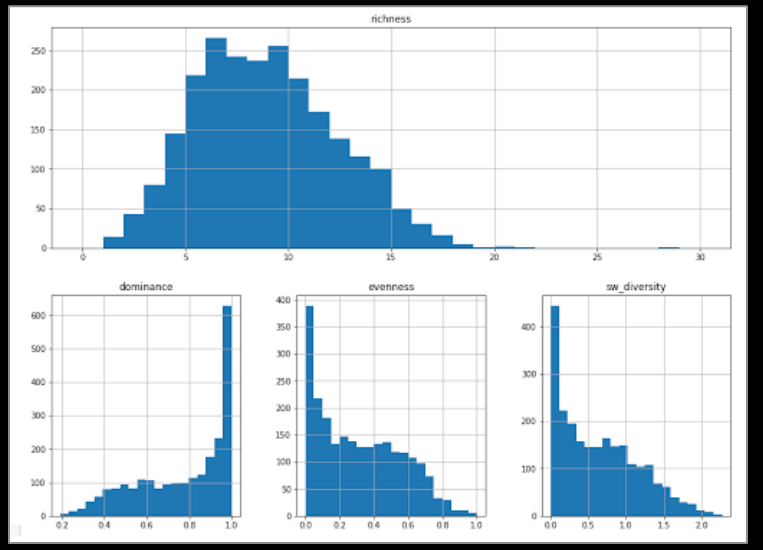
species evenness quantifies how close in count each species is within a sampling event

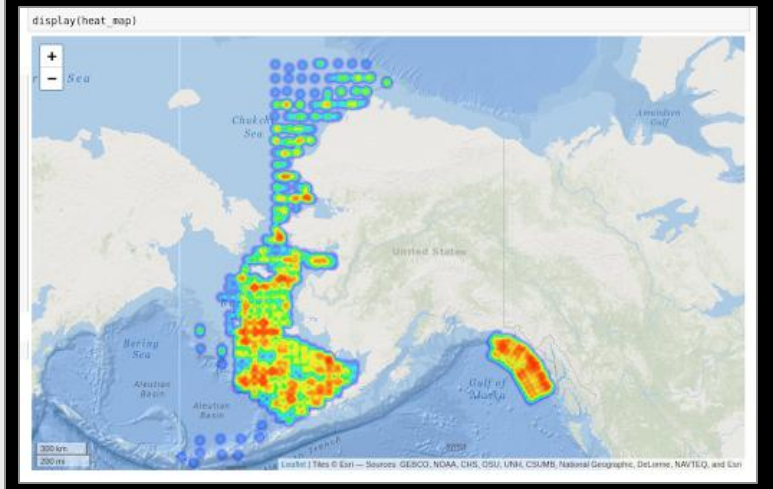
$$J' = \frac{H'}{ln(S)}$$

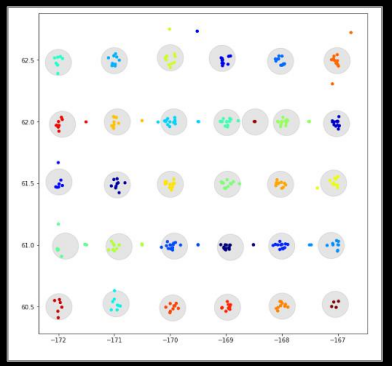
In [17]:
```
# create dominance and shannon-weaver diversity indices
p_i_stats = aggregated_df.groupby(['location_id', dateField])['p_i'].agg
                                                    'sw_diversity':  lamb

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
```

Out[17]:

	location_id	start_date	lat_station	richness	lon_station	sw_diversity	dominance	even
0	1	2002-09-19 18:30:00	51.295745	4	-178.344080	1.036659	0.587719	0.74
1	4	2002-09-12 20:08:00	51.718330	4	-179.724165	0.352634	0.920635	0.25
2	5	2002-09-						







Live Demo

http://mbon.ioos.us/

Axiom DATA SCIENCE