

IOOS Code Sprint Lightning Talks #1

October 8, 2019

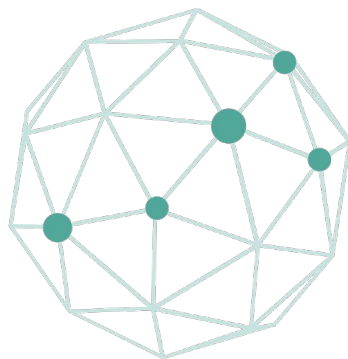


CIOOS update

Ray Brunsting

CIOOS

CANADIAN INTEGRATED OCEAN
OBSERVING SYSTEM



SIOOC

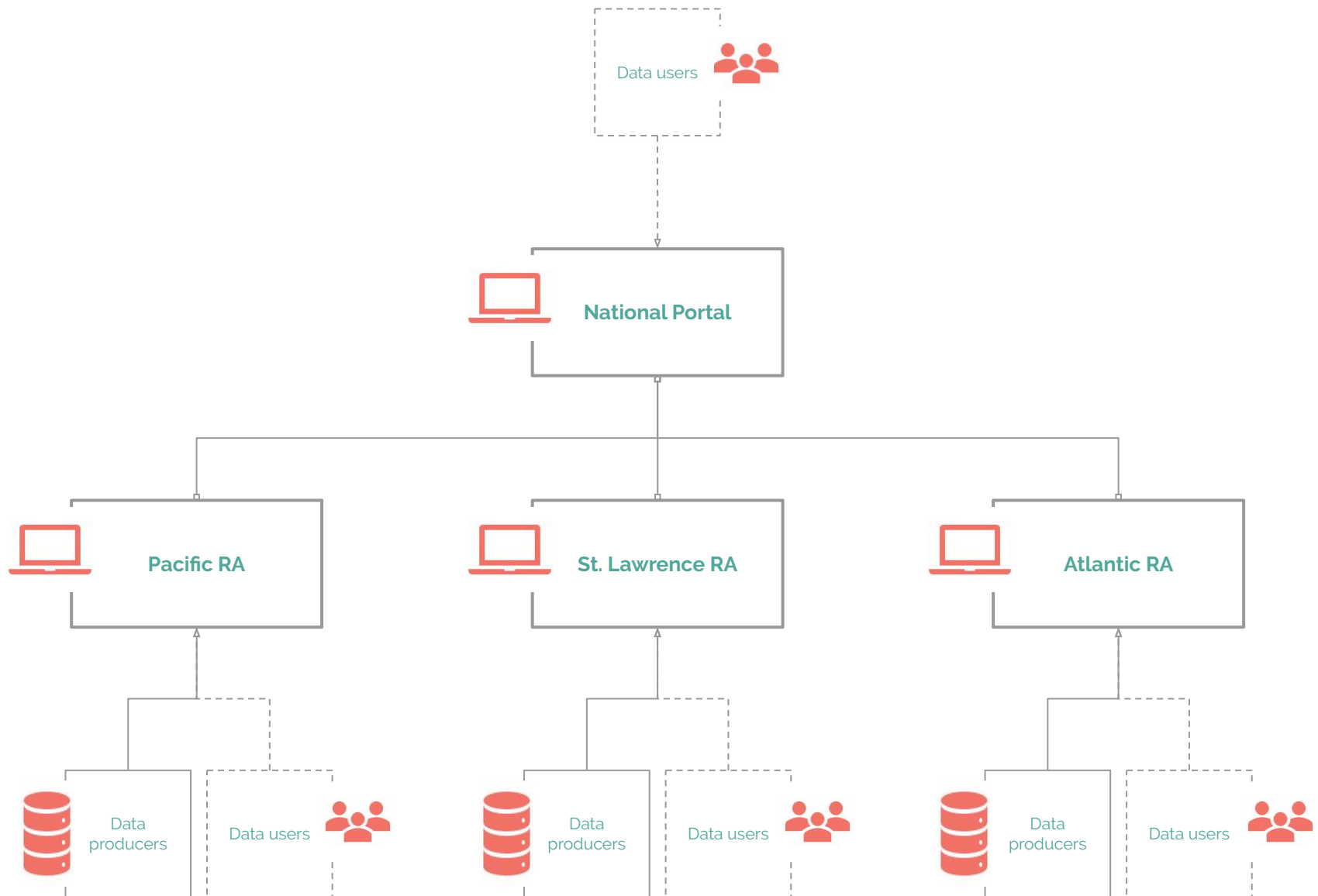
SYSTÈME INTÉGRÉ D'OBSERVATION
DES OCÉANS DU CANADA

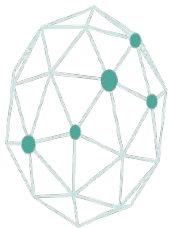
October 8, 2019

CIOOS Update US IOOS 2019 Code Sprint

Ray Brunsting - ray@tula.org

About CIOOS

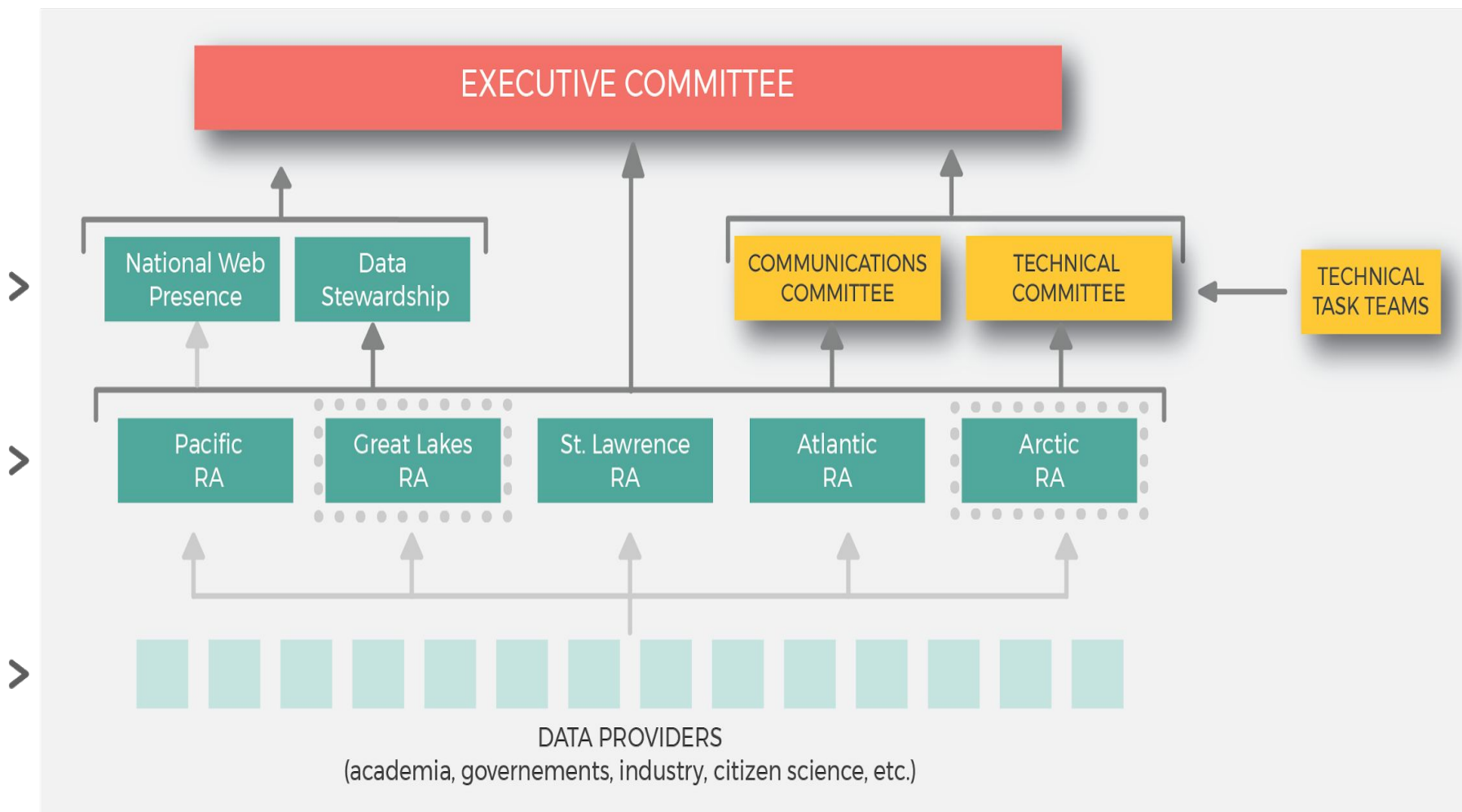




CIOOS

CANADIAN INTEGRATED
OCEAN OBSERVING SYSTEM

Functional / Reporting link ———
Data / Metadata Aggregation link ———
Future Regional Association ● ● ● ● ●





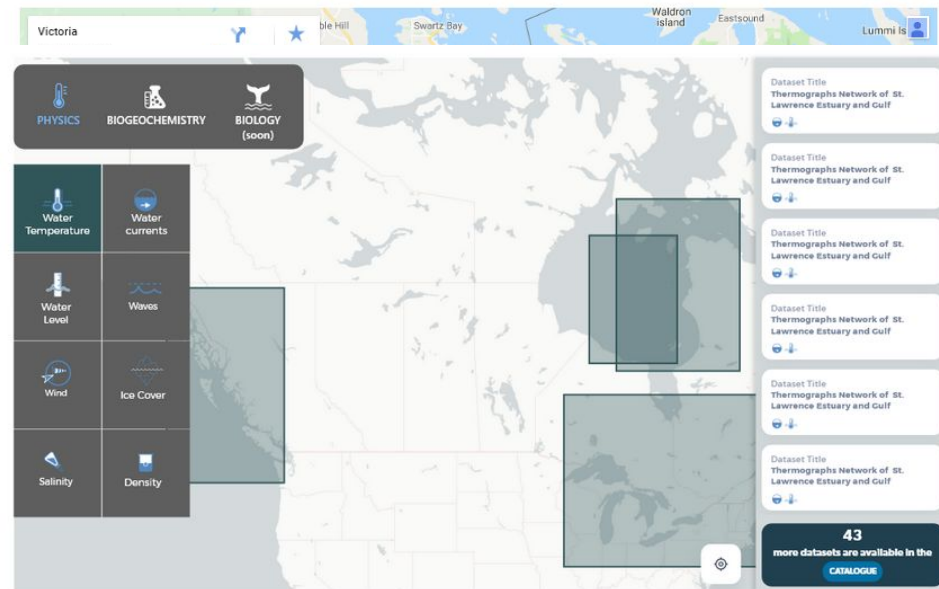
[HOME](#)

[ABOUT](#)

[REGIONS](#)

[DATA TOOLS](#)

[DATA CATALOGUE](#)



sit amet iorem auctum, tristique metus sit amet, egestas neque. integer eros ante, iuctus ac iectus quis, vulputate iooortis orci. vestibulum scelerisque commodo lectus. Phasellus id velit vestibulum, interdum magna id, consequat sem. Suspendisse ornare arcu nec placerat accumsan. In eget eleifend nunc. Duis condimentum diam non tellus bibendum, et eleifend mi malesuada. Mauris urna purus, tristique nec aliquet ac, facilisis a enim.



[EXPLORE DATA](#)

[CONTRIBUTE TO CIOOS](#)

Regional websites



CIOOS | CANADIAN INTEGRATED
OCEAN OBSERVING SYSTEM

REGIONAL ASSOCIATION NAME | FR  

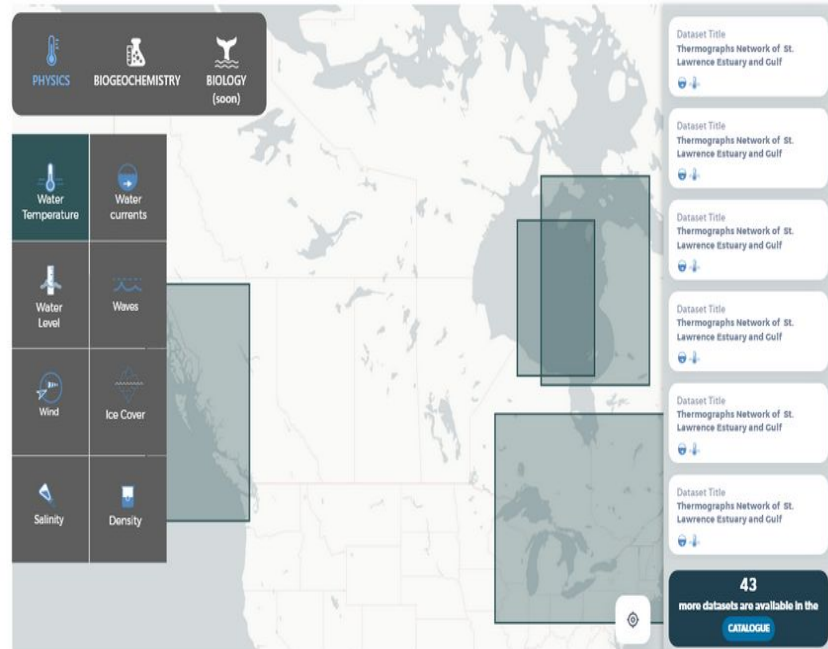
LOGO HERE

[HOME](#)

[ABOUT](#)

[DATA TOOLS](#)

[DATA CATALOGUE](#)



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer gravida sed magna eget pellentesque. Duis pulvinar leo at nunc convallis, et molestie dui suscipit. In sit amet lorem dictum, tristique metus sit amet, egestas neque. Integer eros ante, luctus ac lectus quis, vulputate lobortis orci. Vestibulum scelerisque commodo lectus. Phasellus id velit vestibulum, interdum magna id, consequat sem. Suspendisse ornare arcu nec placerat accumsan. In eget eleifend nunc. Duis condimentum diam non tellus bibendum, et eleifend mi malesuada. Mauris uma purus, tristique nec aliquet ac, facilisis a enim.

[EXPLORE DATA](#)

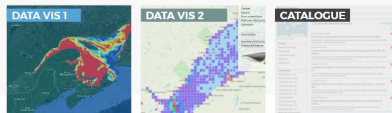
[CONTRIBUTE TO \[RA\]](#)

EXPLORE DATA

National and Regional websites



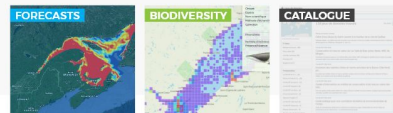
EXPLORE DATA



LATEST NEWS



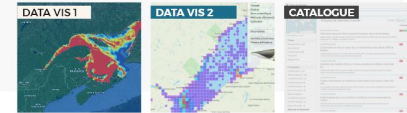
EXPLORE DATA



LATEST NEWS



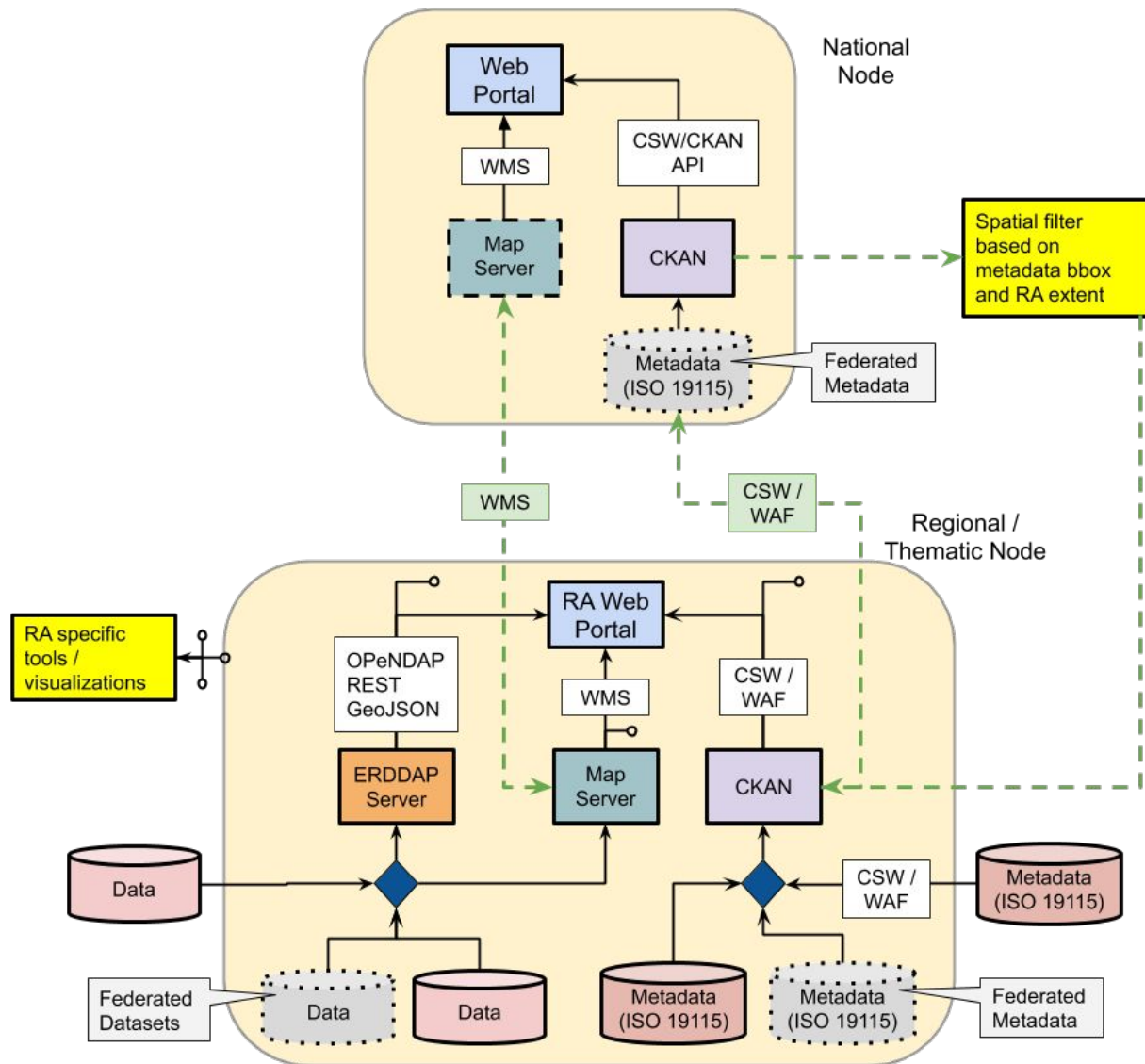
EXPLORE DATA



LATEST NEWS



Solution Architecture



Talk #4 -

Google Charts directly from ERDDAP using .dataTable

Roland Schweitzer

Steps to make a chart

1. Form the URL (either in JS or via Web or both)
2. Download the .dataTable JSON (jQuery or whatever) and instantiate the DataTable object.
3. Configure and draw your chart.

Step 2. - get the data

```
var locationData = $.ajax({  
    url: "ERDDAP URL",  
    dataType: "json",  
    async: false  
}).responseText;
```

Step 2 - instantiate the DataTable

```
var locdata =  
    new google.visualization.DataTable(locationData);
```

Step 3 - configure the chart

```
var moptions = {  
    zoomLevel: 6,  
    displayMode: 'markers',  
    height: 370  
};
```

Step 3 - Draw the chart

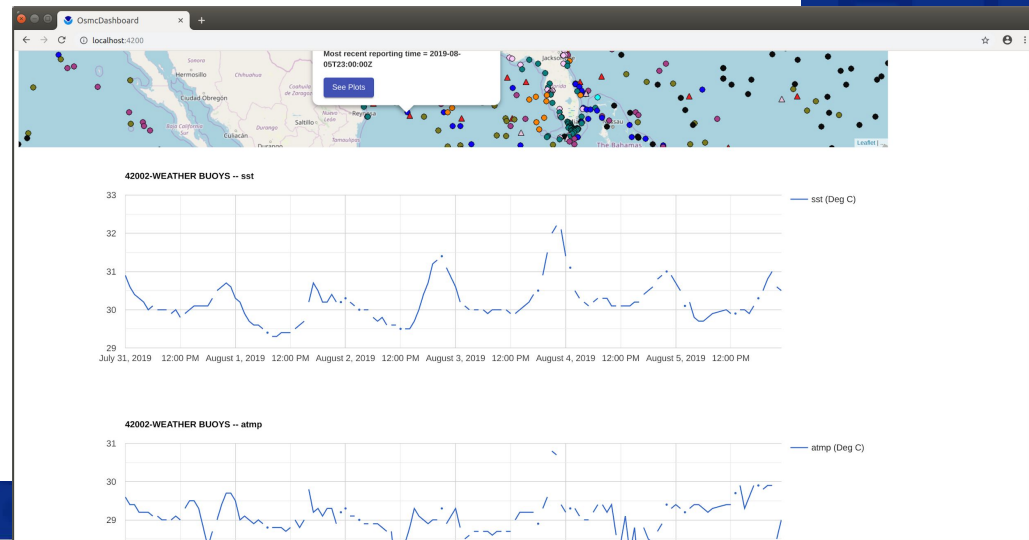
```
var mapchart =  
  new google.visualization.Map(  
    document.getElementById( 'map_div' )  
  );
```



```

this.constrained = this.dataUrl + 'time%2Csst%2Datmp&platform_code=' + this.platformCode + '&orderBy("time")'
return this.httpClient.get<string>(this.constrained).subscribe(value => {
  this.dataTable = new google.visualization.DataTable(value);
  this.atmpChartData = {
    chartType: 'LineChart',
    dataTable: this.dataTable,
    view: {columns: [0,2]},
    options: {
      title: 'Awesome Air Temperature Data Chart',
      explorer: {
        axis: 'horizontal'
      }
    }
  }
  if (this.hasData(this.dataTable, [0,2])) {
    this.has.set('atmp', true);
  }
}

```

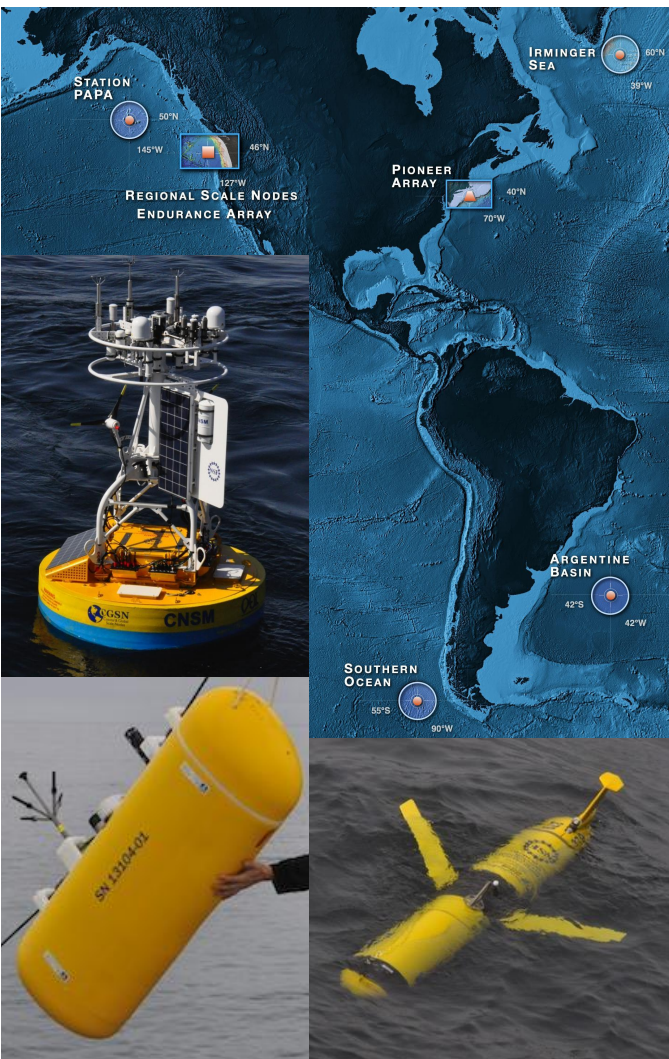
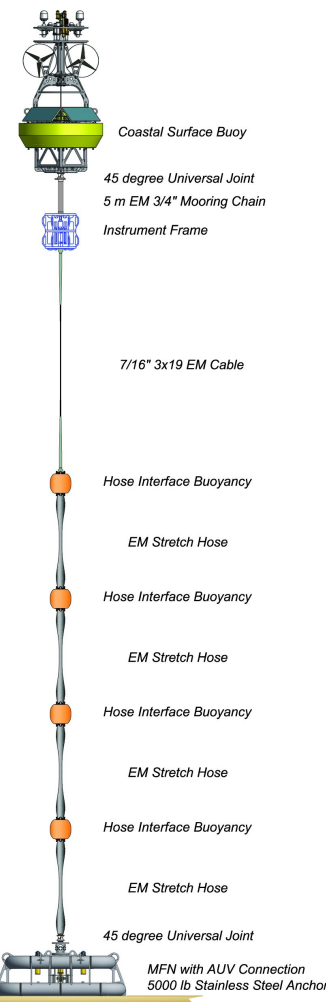


Talk #3 OOI - current status and future goals

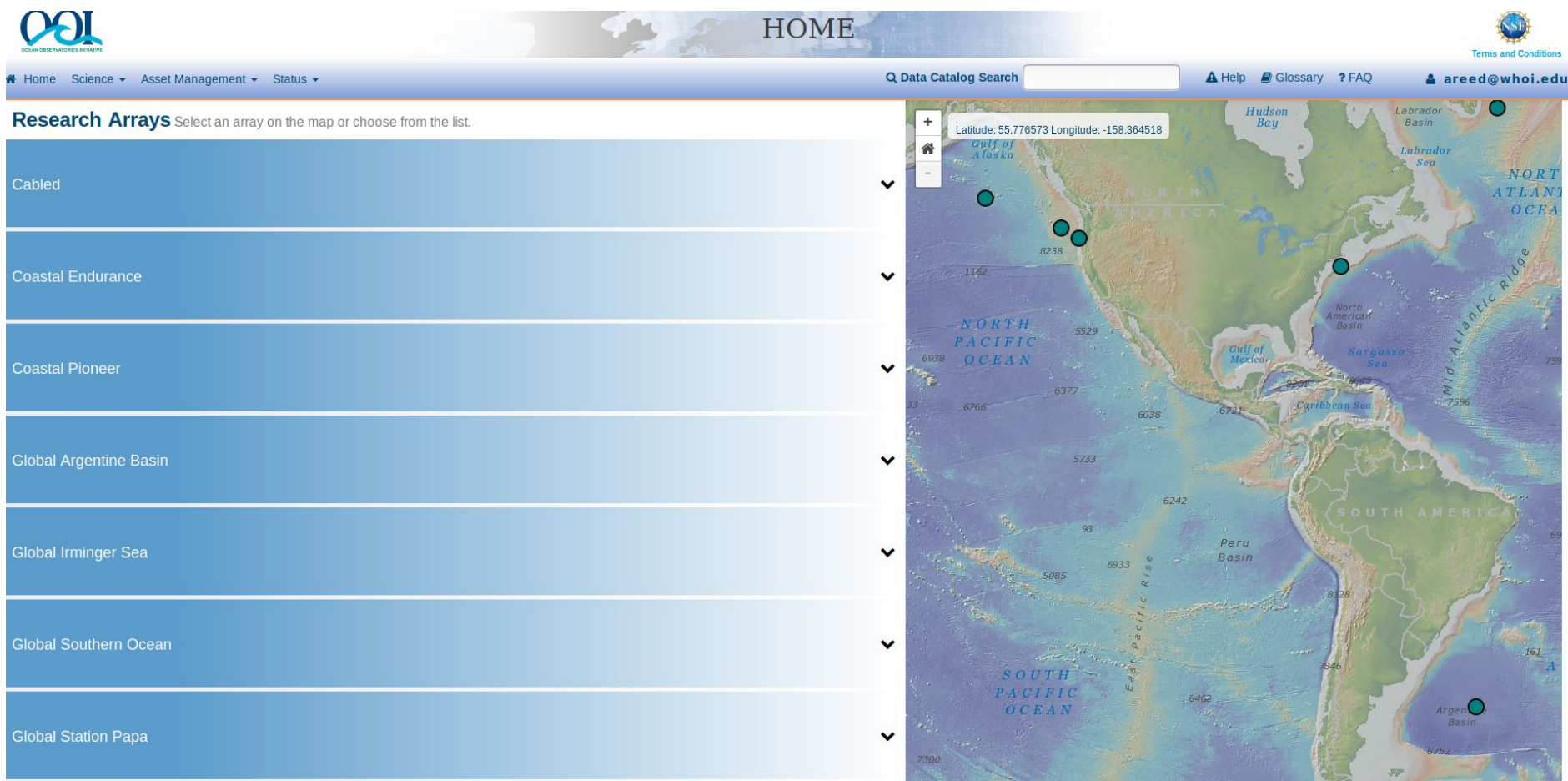
Andrew Reed

Ocean Observatories Initiative

- NSF-funded project for long-term (20+ years) permanent infrastructure for ocean observations
- Two parts:
 - Regional instrumented cabled array on the Oregon-Washington shelf
 - Coastal & Global Scale Nodes (CGSN) with moorings in the Irminger Sea, at Station Papa, in the Southern Ocean & Argentine Basin (descope), and arrays at the mid-Atlantic bight and the Oregon Coast
- CGSN contains:
 - Surface moorings equipped with instruments on the buoy, at 7m depth, and on the anchor
 - Profiling moorings
 - Gliders (3 & 6 month deployments)
 - AUVs



OOI Data – Data Portal (ooinet.oceanobservatories.org)



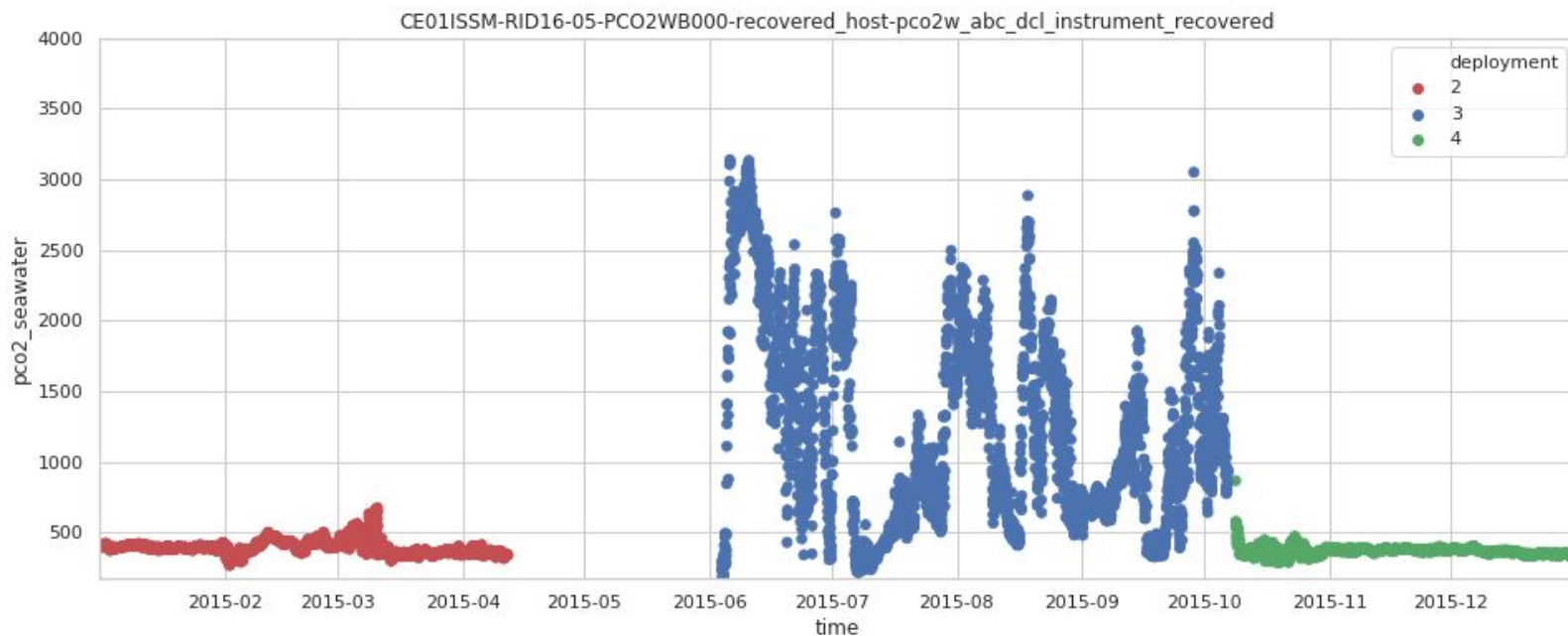
The screenshot displays the OOI Data Portal interface. At the top, there is a navigation bar with the OOI logo, a 'HOME' link, and a search bar labeled 'Data Catalog Search'. Below the navigation bar, a 'Research Arrays' section is visible, featuring a list of research arrays on the left and a map of the North and South Atlantic Oceans on the right. The list of research arrays includes:

- Cabled
- Coastal Endurance
- Coastal Pioneer
- Global Argentine Basin
- Global Irminger Sea
- Global Southern Ocean
- Global Station Papa

The map on the right shows the North Atlantic and South Atlantic Oceans, with various research arrays marked by colored dots. A tooltip is visible over a dot in the North Atlantic, displaying the coordinates: Latitude: 55.776573 Longitude: -158.364518. The map also shows the Gulf of Alaska, Gulf of Mexico, Caribbean Sea, and various basins and ridges.

OOI Data – Quick Overview

- Data may be accessed programmatically via api
- Can be downloaded synchronously or asynchronously (as netCDF) w/CF-vocab
- Data is split based on “streams” which identify: the array, node, instrument, method
 - May be further split based on deployment
 - Overlapping instruments per “stream” requires segregation by deployment
- Current QC
 - No QARTOD Implementation
 - HITL consists of “annotations”: human-readable text based on stream and time



Future Goals

- Currently working on implementing QARTOD QC checks
 - In partnership with Axiom Data Science
 - Moving via test through our data streams on a per-instrument basis
 - Includes nutrient, pH, pCO₂, zooplankton, other biological and chemical data
 - Added functionality needed, particularly with effective dating of test value-ranges
- Making our extensive ship-based water sampling datasets available
 - Using a modified GO-SHIP vocabulary and fields
 - Working with LTER due to co-located stations and sampling plans
- Integrate with the broader Ocean-Data Management and Service community
 - Bring our data formats up to community standards
 - Goal to be able to serve our data through as many portals as possible
 - Make it discoverable through portals such as GOA-ON, partner IOOS orgs (i.e. NaNOOS), project page on BCO-DMO, push to GliderDAC and NBDC
- Value-add data products and services
 - In-depth examination of calibrations, in-situ and co-located data comparisons, etc
 - Provide visualization and data-analysis tools for our users to enhance basic functionality

Talk #5 Using R & Github for Automated Reporting

Ben Best

Automated Reporting for IEA , Sanctuaries

- Data-driven storytelling: interactive infographics
- Real-time management: automated workflows

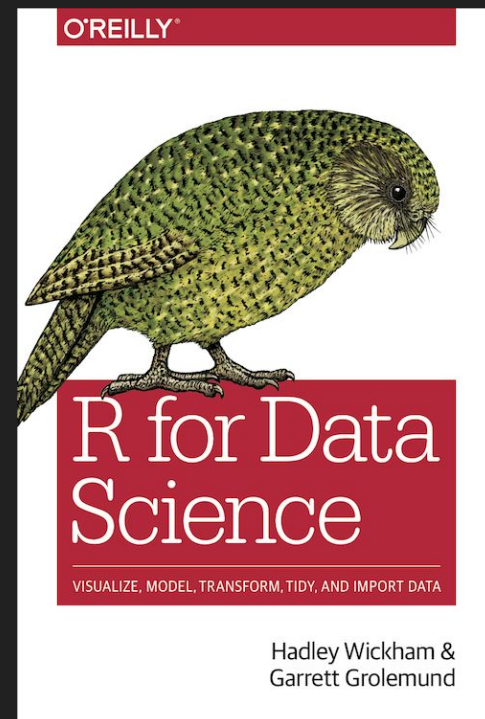
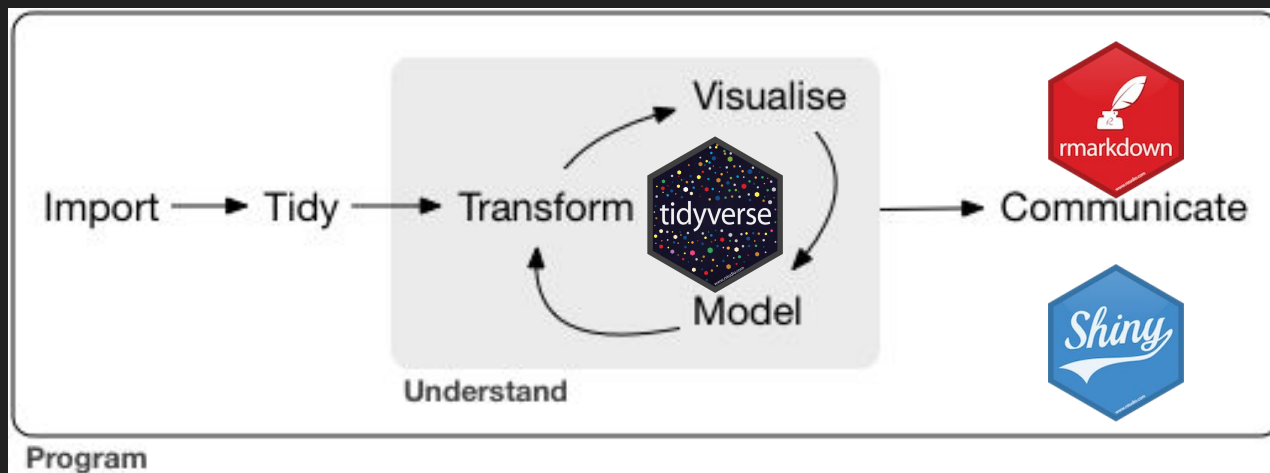
Integrated Ecosystem Assessment (IEA) Ecosystem Status Reports (ESRs) x 1 yr



National Marine Sanctuaries (NMS) Condition Reports (CRs) x 5-7 yrs



R for Data Science



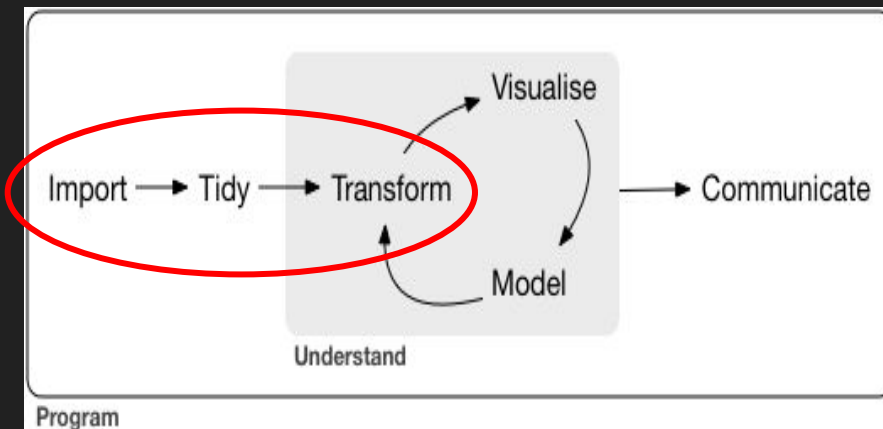
Exploratory Data Analysis

- book: Tukey (1977) Exploratory Data Analysis
- software: **S** (at Bell Labs) → **S-plus** → **R**

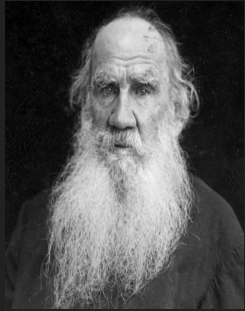
For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights

nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html

“Data scientists, according to interviews and expert estimates, spend from **50 to 80 percent** of their time mired in the mundane labor of **collecting and preparing data**, before it can be explored for useful information.” — NY Times (2014)



“Happy families are all alike; every unhappy family is unhappy in its own way.”
— Leo Tolstoy



“Tidy datasets are all alike, but every messy dataset is messy in its own way.”
— Hadley Wickham



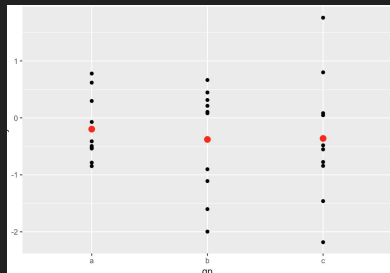
Visualization

Static: [ggplot2](#)

“grammar of graphics”

```
ggp: data aes aesthetics
geom_point() +
geom_point(
  data = ds,
  aes(y = mean),
  color = 'red', size = 3)
```

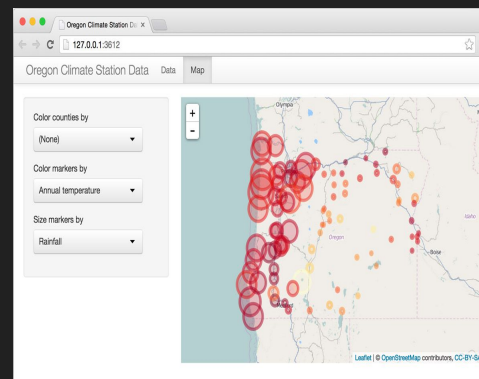
geometries



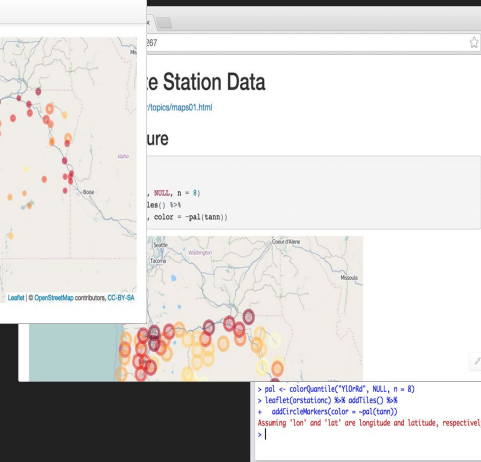
Interactive: [htmlwidgets](#)

```
library(leaflet)
pal <- colorQuantile("YlOrRd", NULL, n = 8)
leaflet(orstationc) %>%
  addTiles() %>%
  addCircleMarkers(color = ~pal(tann))
```

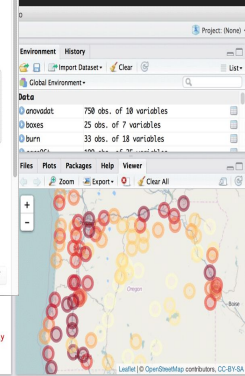
Shiny



Rmarkdown

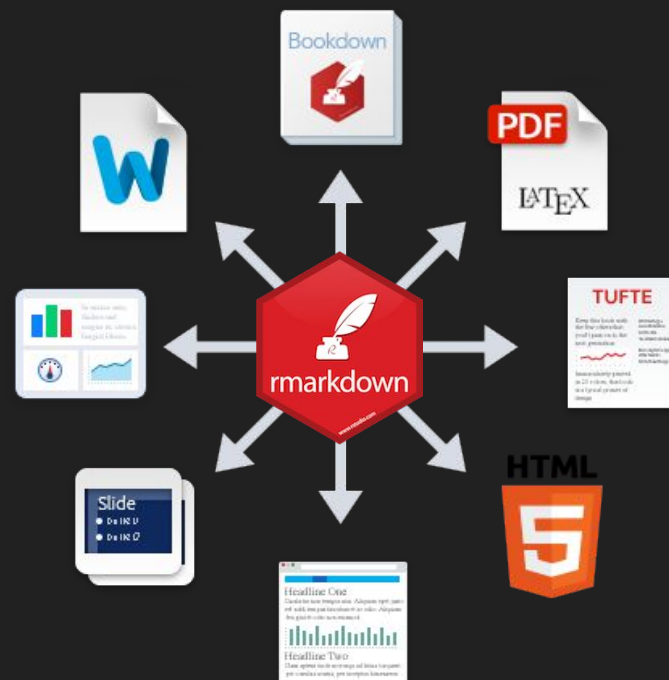


RStudio



Reproducible Research with R and RStudio

Root			
Research Project			
Data	Analysis	Presentation	
MainData.csv	MainAnalysis.R	Article	Other
GatherSource	ResultsFigures	Paper.Rnw	Slideshow
Makefile	Figure1.R	Main.bib	Slideshow.Rnw
MergeData.R	Figure2.R	Packages.bib	Website
Gather1.R	Figure3.R	figure	Website.Rmd
Gather2.R		Figure1.pdf	
Gather3.R		Figure2.pdf	
		Figure3.pdf	



1min video: [What is Rmarkdown?](#)

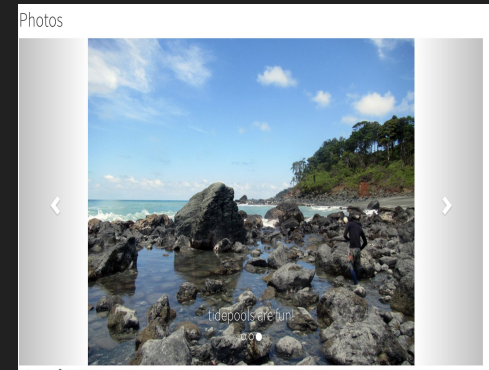
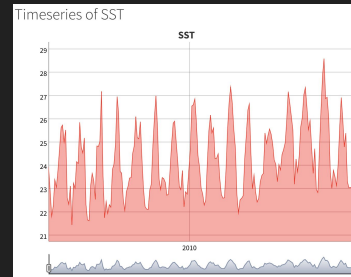
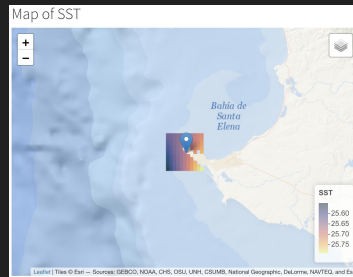
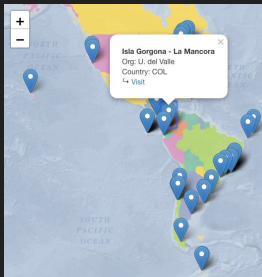
uses: [pandoc](#)

MBON Pole to Pole

“Community of practice” for rocky intertidal + sandy beach across the Americas

Home

Site



marinebon.github.io/p2p

Brazil workshop 2018-08



Ecosystem Infographics

tech: **d3 JS**



IEA produces Ecosystem Status Reports annually with ecosystem elements
Goal: make ecosystem conceptual diagrams interactive and supported with maps & time series of data

Overview



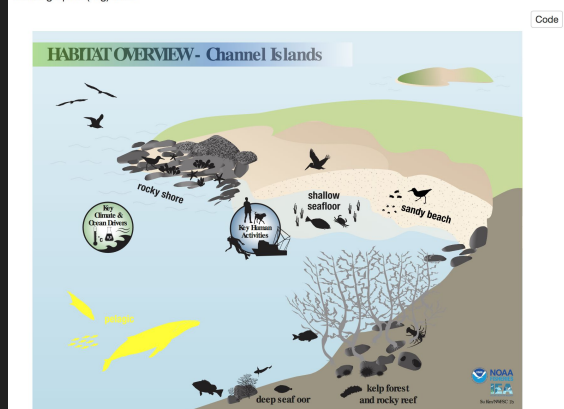
Habitat



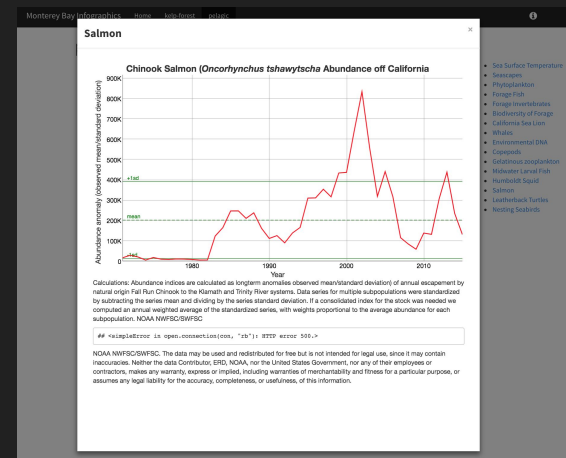
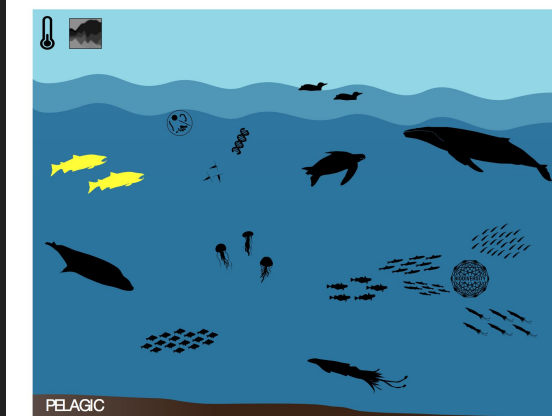
Data Popup

CINMS Ecosystems Overview

Example of using `r2d3` over previously custom built `marinebon/infographiq` for creating interactivity in scalable vector graphic (svg) files.



Pelagic

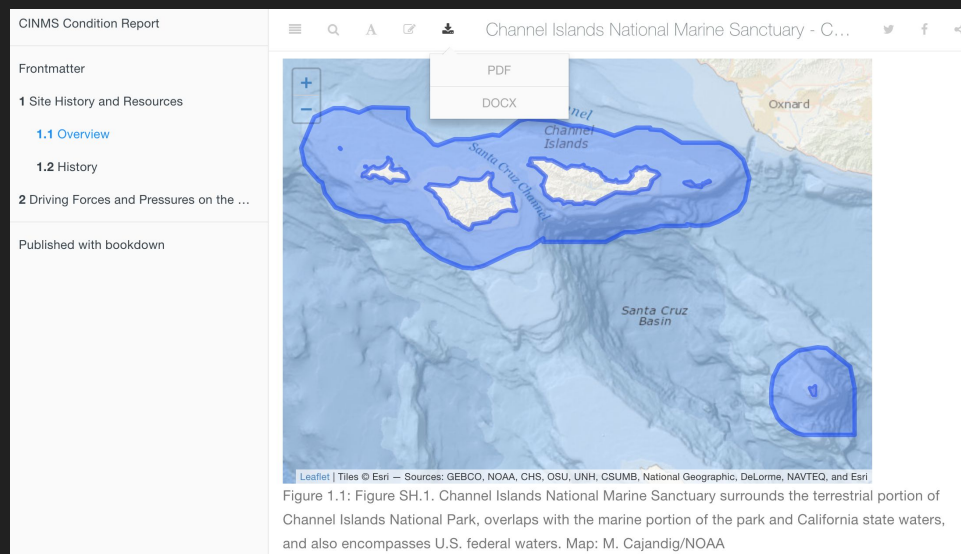
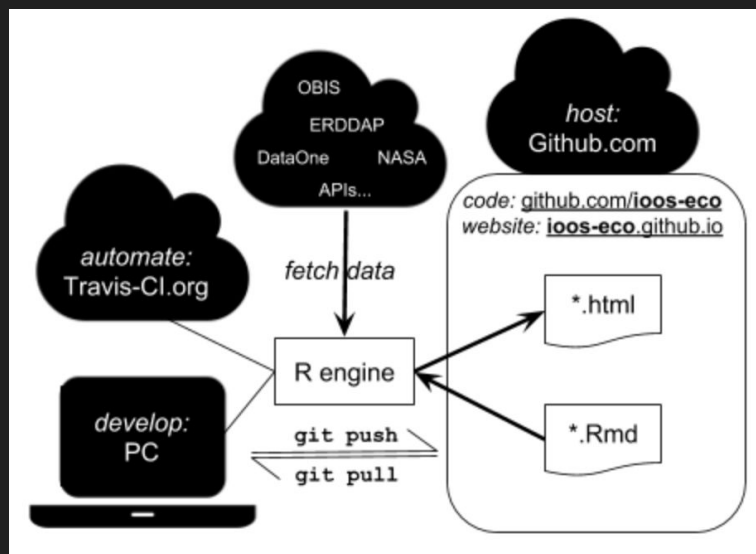


Sanctuary Condition

tech: **R** **bookdown**



NOAA National Marine Sanctuaries produce Condition Report every 5-7 years
Goal: make condition reports online, interactive and up to date



marinebon.github.io/cinms-cr ([code](#))

demo: r4ioos

bbest.github.com/r4ioos

- index.html

r4ioos

demo of R for automation & reporting

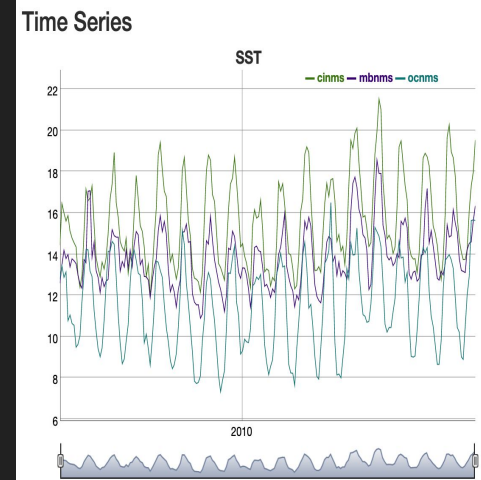
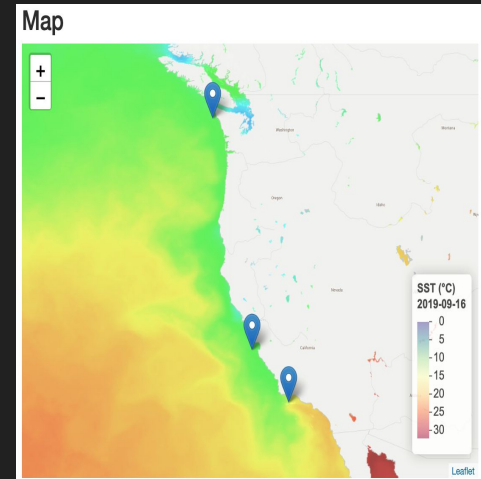
build passing

- test.html
- test.docx
- test.pdf

- make.R

```
library(rmarkdown)
render("README.md", output_format = "html_document",
      output_file = "index.html")
render("test.Rmd", output_format = "html_document")
#render("test.Rmd", output_format="pdf_document")
#render("test.Rmd", output_format="word_document")
```

- test.html



.travis.yml

github.com/bbest/r4ioos

```
language: r
dist: xenial
latex: false

cache:
  packages: true
  directories:
    - $HOME/bin

script:
  - Rscript make.R > make.log

deploy:
  provider: pages
  skip_cleanup: true
  github_token: $GITHUB_TOKEN
  on:
    branch: master
  local_dir: .
```

```
# Linux dependencies: rerddap, leaflet
addons:
  apt:
    sources:
      - sourceline: 'ppa:ubuntugis/ppa'
    packages:
      - libudunits2-dev
      - libproj-dev
      - libgeos-dev
      - libgdal-dev
      - netcdf-bin

notifications:
  email:
    recipients:
      - ben@ecoquants.com
  on_success: change
  on_failure: change
```

details:
[GitHub Pages Deployment
dpl-v1 - Travis CI](#)

future:
[dpl-v2 - Travis CI](#)

Installing R packages into Travis-CI environment

- DESCRIPTION

```
Package: placeholder
Title: Does not matter.
Version: 0.0.1
Imports: rmarkdown,
  dplyr, dygraphs, here, glue, leaflet, lubridate,
  readr, rerddap, purrr, stringr, tidyr,
  xts
Remotes: ropensci/rerddap
```

Travis-CI settings for repo

- authenticate

Environment Variables

Customize your build using environment variables. For secure tips on generating private keys [read our documentation](#)

GITHUB_TOKEN



.....

Available to all branches

- run daily/weekly/monthly

Cron Jobs

master

Runs daily

Ran -

Scheduled in about a min Do not run if there has been a recent commit



And now for something
completely different ...

Whirlyball, anyone?

