

# **Hacking the Academy: Open in Action**

## OA Week @ University of Washington

---

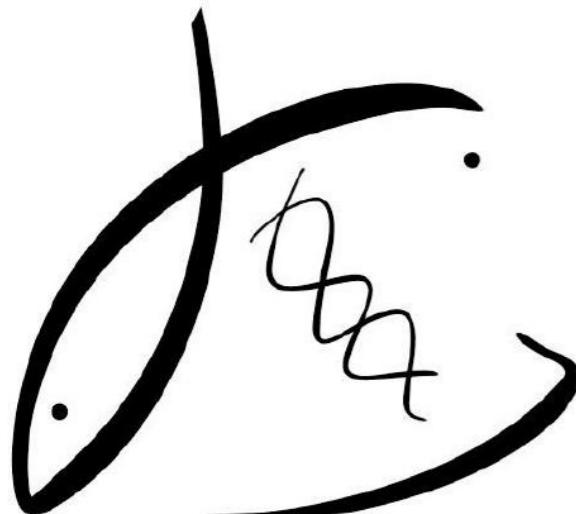
Steven Roberts

Kenneth K. Chew Endowed Professor  
School of Aquatic and Fishery Sciences

University of Washington

[robertslab.info](http://robertslab.info)

@sr320



# Open Access Week

---

- You are free to Share!

These slides plus links etc @  
[github.com/sr320/talk-UW-OA-2016](https://github.com/sr320/talk-UW-OA-2016)

# Outline

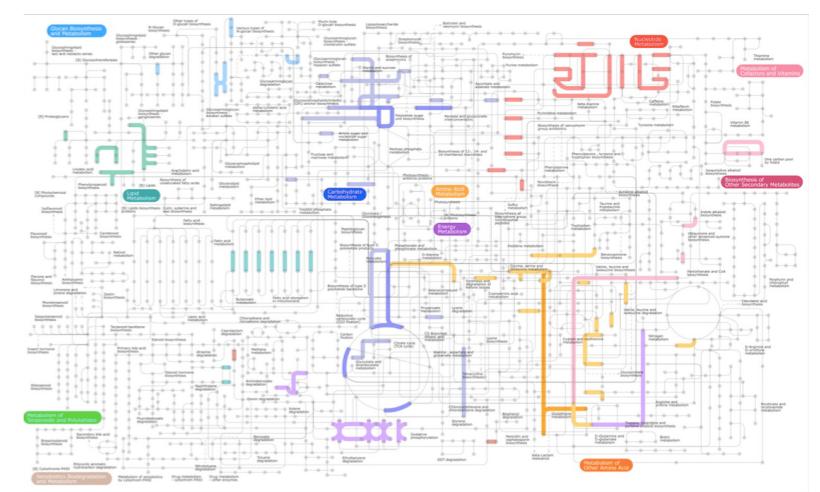
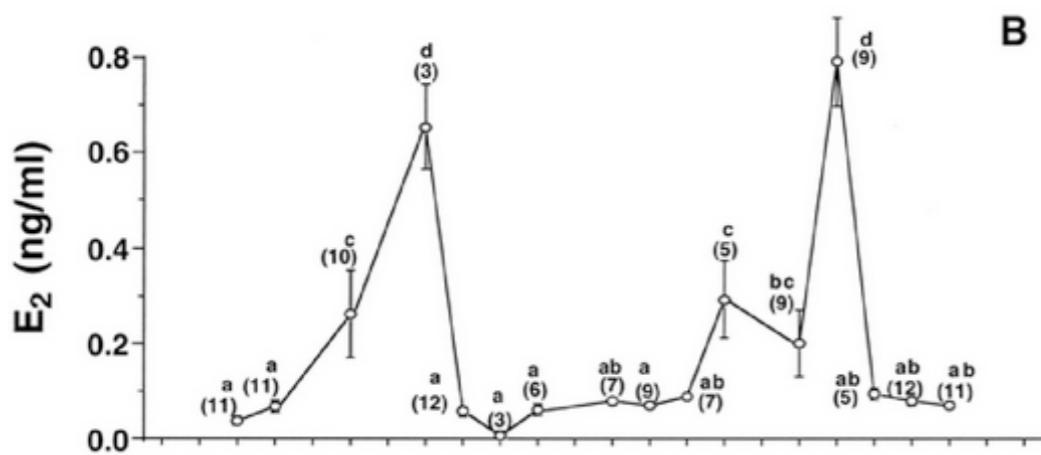
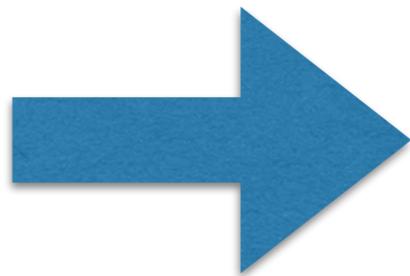
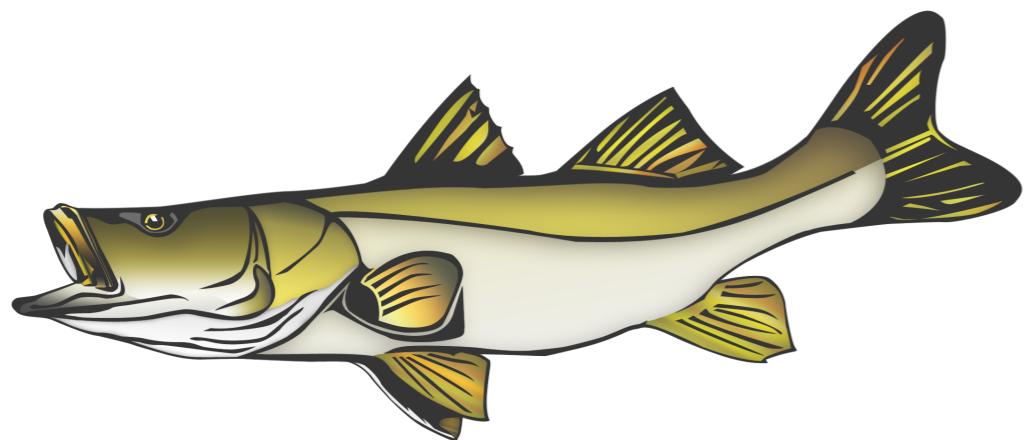
---

- Context
- **What** is Open and Online
- **Why** Open and Online

# Context

## Physiology

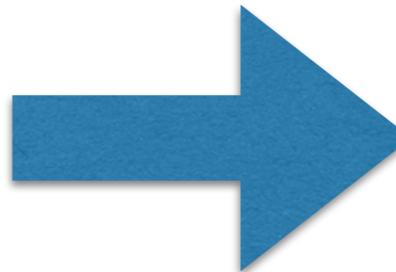
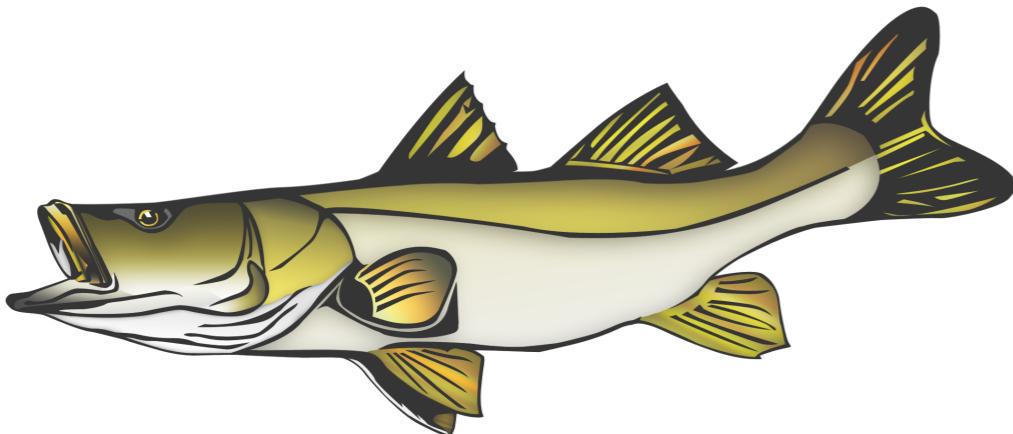
*How fundamental processes work in aquatic species*



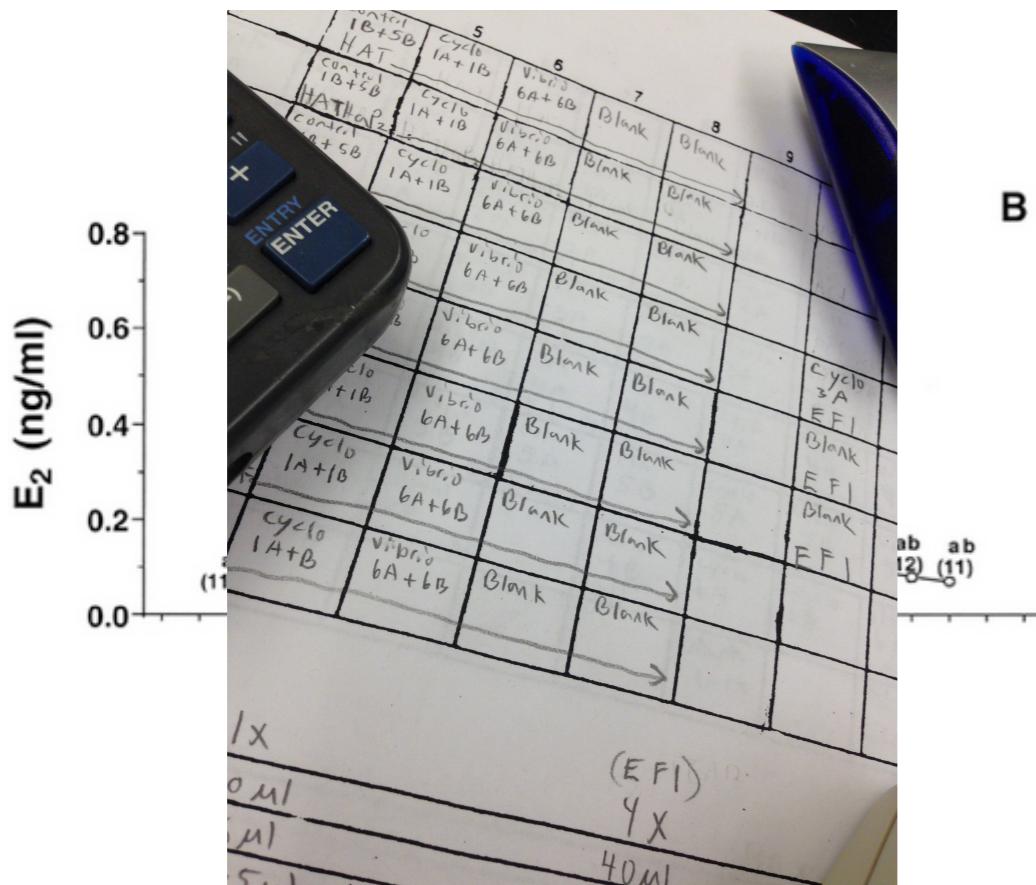
# Context

## Physiology

*How fundamental processes work in aquatic species*



B

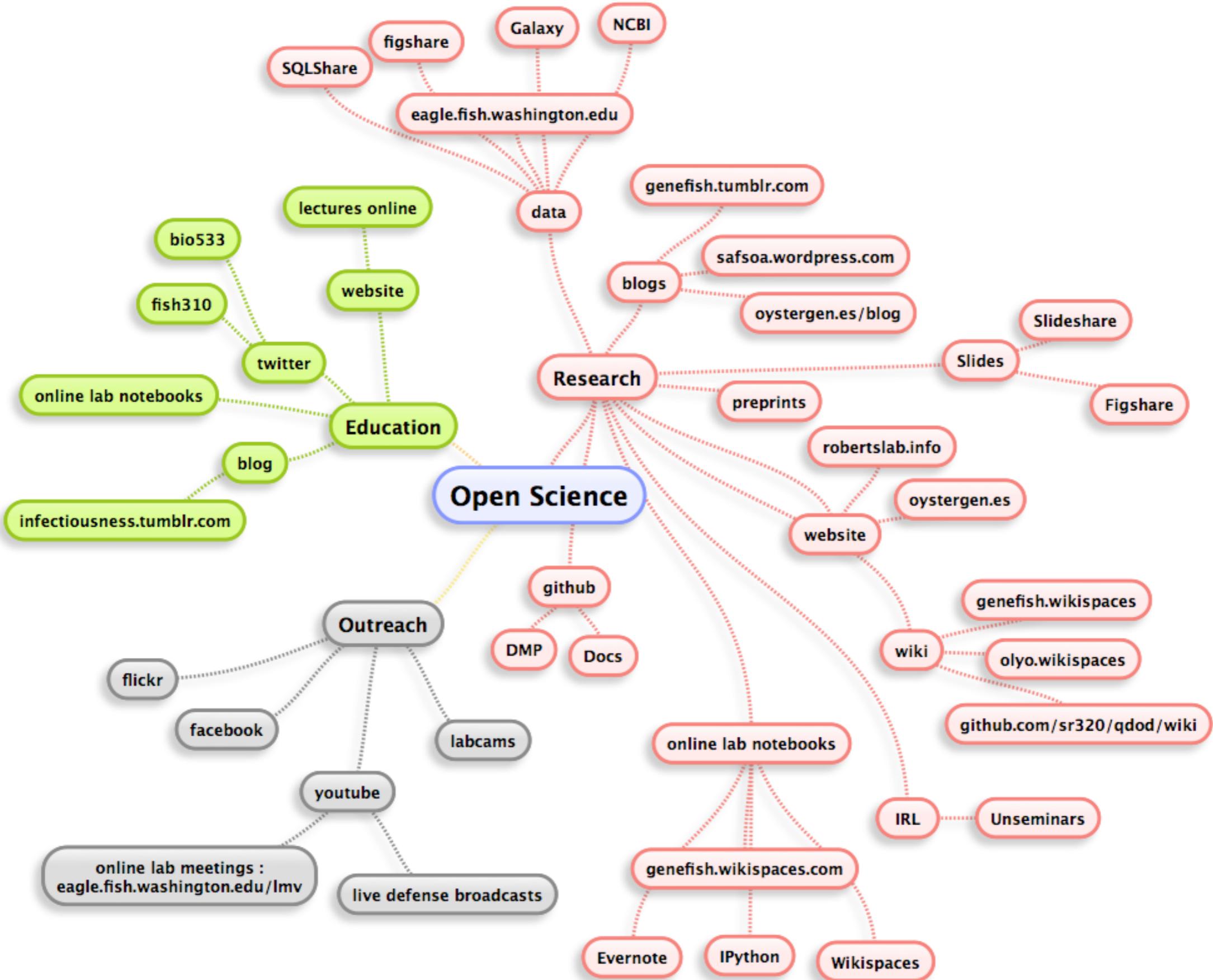


# Data Rich

*Data is  
Open &  
Online*

*Data is  
Open &  
Online*

***So is everything else***



# Approach



Data Lake

# Data Acquisition and Analysis

---

“But are you not worried that someone is going to ‘scoop’ your data?”

# Data Rich

Everyone has backlogs of data that is not yet analyzed, no one is sitting around with nothing else to do

# Data Acquisition and Analysis

## Open Notebook Science

“When you’re working in the open, only 1 person in the ENTIRE WORLD needs to solve a given problem FOREVER”  
- Jean-Claude Bradley

TWEETS 3,238 FOLLOWING 703 FOLLOWERS 868 LIKES 1,006

**OpenNotebookScience** @ONScience FOLLOWS YOU  
Updates from #openscience notebooks, brought to you by @openscience. Start yours at [onsnetwork.org/joinons/](http://onsnetwork.org/joinons/)  
[onsnetwork.org](http://onsnetwork.org)  
Joined October 2013

[Tweet to](#) [Message](#)

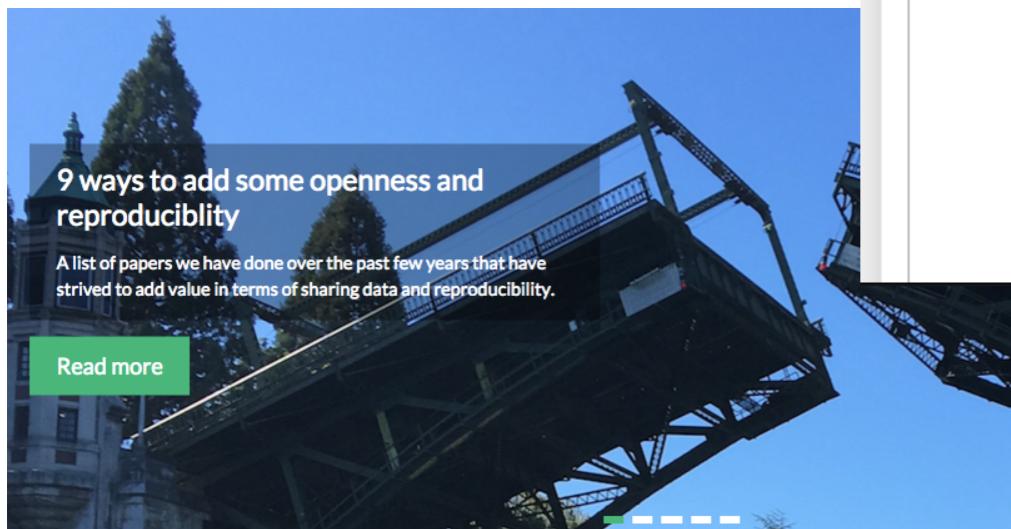
**Tweets** [Tweets & replies](#) [Media](#)

Pinned Tweet  
**OpenNotebookScience** @ONScience · 3 Aug 2014  
Really impressed with the open notebooks being kept by @sr320's @EIMD\_fhl students [onsnetwork.org/eimd/](http://onsnetwork.org/eimd/) #openscience as a matter of course

[...](#)

# Data Acquisition and Analysis

# Open Notebook Science



jupyter Welcome3 (autosaved)

File Edit View Insert Cell Kernel Help

Python 3

Cell Toolbar: None

Welcome

Welcome to Pineapple, the next generation scientific notebook.

Run Python code

In [1]: `2 ** 64`

Out[1]: 18446744073709551616

You can make plots right in the notebook

In [2]: `%matplotlib inline  
import matplotlib.pyplot as plt  
import numpy as np  
x = np.linspace(0, 10)  
plt.plot(x, np.sin(x));`

jldimond / Coral-CpG

Unwatch 2

Repository associated with article "Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change"

22 commits 1 branch 0 releases 2 contributors

Branch: master +

Update READMEs

jldimond authored 5 days ago latest commit d94c7b5aec

analyses Update READMEs 5 days ago

data typo 7 days ago

ipynb Update READMEs 5 days ago

README.md typo 7 days ago

README.md

Analyses and data associated with manuscript:  
*Germline DNA methylation in six species of reef corals: patterns and potential roles in response to environmental change*

A GitHub repository page for 'Coral-CpG' by jldimond. The repository is associated with a manuscript about germline DNA methylation in reef corals. It shows 22 commits, 1 branch, 0 releases, and 2 contributors. The 'master' branch is selected. A list of recent commits includes updates to READMEs, analyses, data, ipynb files, and the README.md file. The repository is described as being associated with a manuscript titled 'Germline DNA methylation in six species of reef corals: patterns and potential roles in response to environmental change'.

Code Issues 20 Pull requests 0 Projects 1 Wiki Pulse

Filters  Labels Milestones

20 Open ✓ 280 Closed

Develop Schematic of Entirety of Oly Oyster Samples

#317 opened an hour ago by sr320

Proteomics Meeting - Developing Timeline

#314 opened a day ago by sr320

Hyak node

#313 opened a day ago by sr320

Take action on genomes..

#310 opened 2 days ago by sr320

System update available for Synology servers. enhancement

#308 opened 2 days ago by kubu4

R Studio Server broken Docker enhancement low priority

#307 opened 2 days ago by kubu4

Do we modify all xserves for use? labmeeting

#299 opened 8 days ago by sr320

Plan to handle extended power outage?

#295 opened 14 days ago by kubu4

Why does a post not break to read more with Jekyll-now

#292 opened 15 days ago by sr320

Consider how to best get tag data online

#288 opened 18 days ago by sr320

Proposal for Data Incubator? labmeeting

#285 opened 22 days ago by sr320

Deploy Shellfish out to BBC labmeeting

#277 opened on Sep 26 by sr320

High School teacher involvement SeaGrant labmeeting

#276 opened on Sep 26 by sr320



[Code](#)[Issues 6](#)[Pull requests 0](#)[Projects 0](#)[Wiki](#)[Pulse](#)[Graphs](#)[Settings](#)[Edit](#)[New Page](#)

# Home

Steven Roberts edited this page 2 days ago · 28 revisions

This course will teach core computing skills as well as project specific approaches. Each student will be developing and completing a research project targeting journal article submission by the end of the Quarter. There will be an emphasis on developing habits that increase automation which in turn will facilitate reproducibility. The course primary course platform will be GitHub, with each student creating their own repository.

T 3:00-4:20 FSH 213  
Th 9:30-11:20 FSH 213

Week	Description	Reading	Quiz
zero	Biology, Course Framework, Getting set-up	Preface xiii-xxv; How to Learn Bioinformatics 1-18	<a href="#">Questions</a>
one	Bash, version control, Project Set-up	Setting Up and Managing a Bioinformatics Project 21-35; Remedial Unix Shell 37-54	<a href="#">Questions</a>
two	Jupyter, Annotation	Retrieving Bioinformatics Data 109-124, Unix Data tools 125-168	<a href="#">Questions</a>
three	Projects	Working with Sequence Data 339-354	<a href="#">Questions</a>

## Pages 3

### Week 3

- [Notes](#)
- [Quiz](#)
- [Project](#)

### Week 4

- [Notes](#)
- [Quiz](#)
- [Project](#)

## Grades

## Class Repos

- [hputnam](#)
- [vaaminiv](#)



Slideshare

Slides

Figshare

genefish.wikispaces

olyo.wikispaces

github.com/sr320/qdod/wiki

wiki

IRL

Unseminars

## online lab notebooks

## genefish.wikispaces.com

Evernote

IPython

Wikispaces

## Outreach

flickr

facebook

youtube

DMP

Docs

labcams

online lab meetings :  
[eagle.fish.washington.edu/lmv](http://eagle.fish.washington.edu/lmv)

live defense broadcasts

Branch: master [master](#) | [Compare](#) | [Pull requests](#)

[LabDocs / Proposal\\_database.md](#)

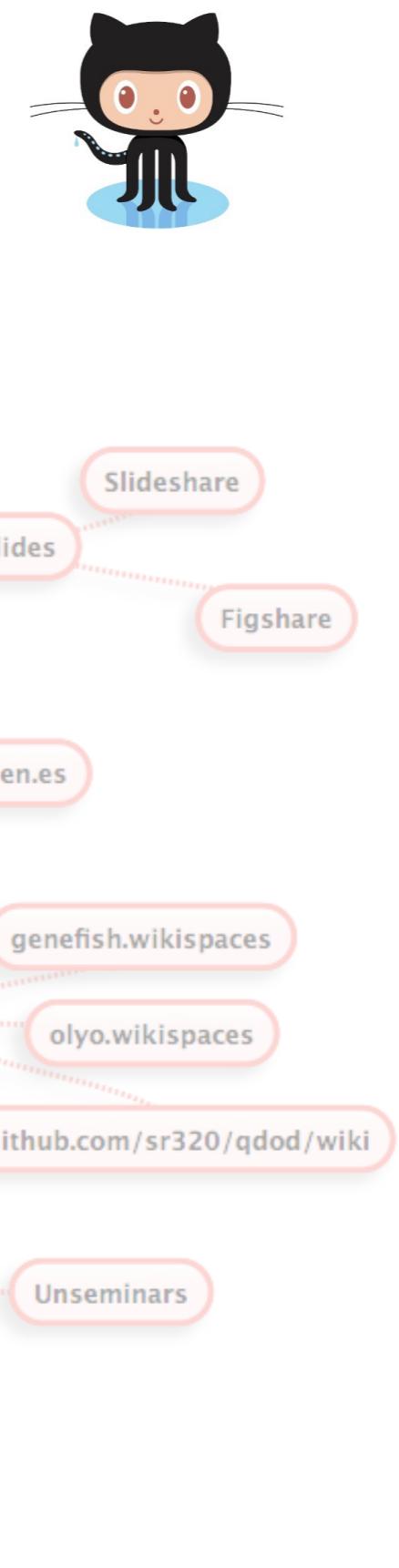
 [sr320](#) dmp

1 contributor

36 lines (35 sloc) | 9.99 KB

[Raw](#) [Blame](#) [History](#)

Agency	Year	Title	Funded (y/n)	Linkout
EPA	2009	Integrated Approach to Assess Emerging Threats to Puget Sound Ecosystem Health	n	<a href="#">PDF</a>
EPA	2008		y	<a href="#">PDF</a>
EVOS	2009	High Density DNA Sequencing to Detect Population Structure of Pacific Herring	y	<a href="#">PDF</a>
JGI	2008	Genome Sequencing of <i>Cycloclasticus pugetii</i> , a Widely Distributed Marine PAH-Degrading Bacterium	y	<a href="#">PDF</a>
NIH	2007	Searching for Olfactory Receptor Genes	n	<a href="#">PDF</a>
NOAA	2013	Using Shotgun Proteomics to Obtain Hemolymph Markers for Determining Sex and Stage of Gonad Maturation in Oysters	-	<a href="#">PDF</a>
NOAA	2007	Sea scallops as sentinels of deepwater pollution	n	<a href="#">PDF</a>
NOAA Aquaculture	2013	Alleviating Regulatory Impediments to Native Shellfish Aquaculture	y	<a href="#">PDF</a>
NOAA ODRP	2007	Oyster Herpes Virus: Understanding Differential Susceptibility, Transmission Dynamics and Host-Pathogen Interactions	n	<a href="#">PDF</a>



**YouTube**

Search

Home Videos Playlists Channels Discussion About

 **Lab Meetings**  
Steven Roberts • 98 videos • 141 views • Last updated on Oct 4, 2016  
Add a description  
▶ Play all Share Playlist settings

This playlist is automatically sorted by date published.

Rank	Video Title	Uploader
1	TWIF 118: Lab Meeting - Sept 26	by Steven Roberts
2	TWIF 117: Lab Meeting - September 8	by Steven Roberts
3	TWIF 116: Lab Meeting - Aug 25	by Steven Roberts
4	TWIF 115: Lab Meeting - August 11	by Steven Roberts
5	TWIF 111: Lab Meeting - June 29	by Steven Roberts
6	TWIF 110: Lab Meeting - June 8	by Steven Roberts



# Why?

“Why would you do this? It has to take so much time and I don’t have that kind of time.”

# Why?

It is easier for *me* to find stuff.

# Approach



Data Lake

# Why?

## UW TODAY

October 7, 2015

### Student collaboration leads to first results describing sick sea star immune response

Michelle Ma

News and Information

"I learned so much about open science and the idea of having open lab notebooks, where everything is available. That has had the biggest impact on how I do my science," said co-author [Sarah Gignoux-Wolfsohn](#), a doctoral student at Northeastern University. "Open science was really integral to us being able to work on this paper together."

*Students.*

Things will change, and there will be no open science.

# Why?

other benefits

# Outline

---

- Context
- **What** is Open and Online
- **Why** Open and Online

“This all sounds fun, but a little overwhelming.”

“This all sounds fun, but a little overwhelming.”

## **Simple tips to start...**

When publishing paper, publish data.



GitHub.

Jupyter.



“This all sounds fun, but a little overwhelming.”

## **Simple tips to start...**

When publishing paper, publish data.



GitHub.

Jupyter.



These slides plus links etc @  
**[github.com/sr320/talk-UW-OA-2016](https://github.com/sr320/talk-UW-OA-2016)**