



DOING SCIENCE : TIPS AND TRICKS

Steven Roberts - University of Washington

I NEED A SLIDEDECK FOR THIS AFTERNOON... WHERE SHOULD I START?

I NEED A SLIDEDECK FOR THIS AFTERNOON... WHERE SHOULD I START?

Screenshot of a GitHub profile page for Steven Roberts (sr320). The page shows pinned repositories, a contribution heatmap, and social links.

Pinned Repositories:

- nb-2020** (Jupyter Notebook)
- RobertsLab/resources** (HTML, 3 stars, 3 forks)
- course-fish546-2018** (Pep8, 3 stars, 2 forks)
- hputnam/EastOyEpi** (Jupyter Notebook, 2 stars)
- course-fish497-2019** (HTML)

Contributions Heatmap: 954 contributions in the last year. The heatmap shows activity levels from 2015 to 2020, with darker shades indicating higher contribution density.

Profile Information:

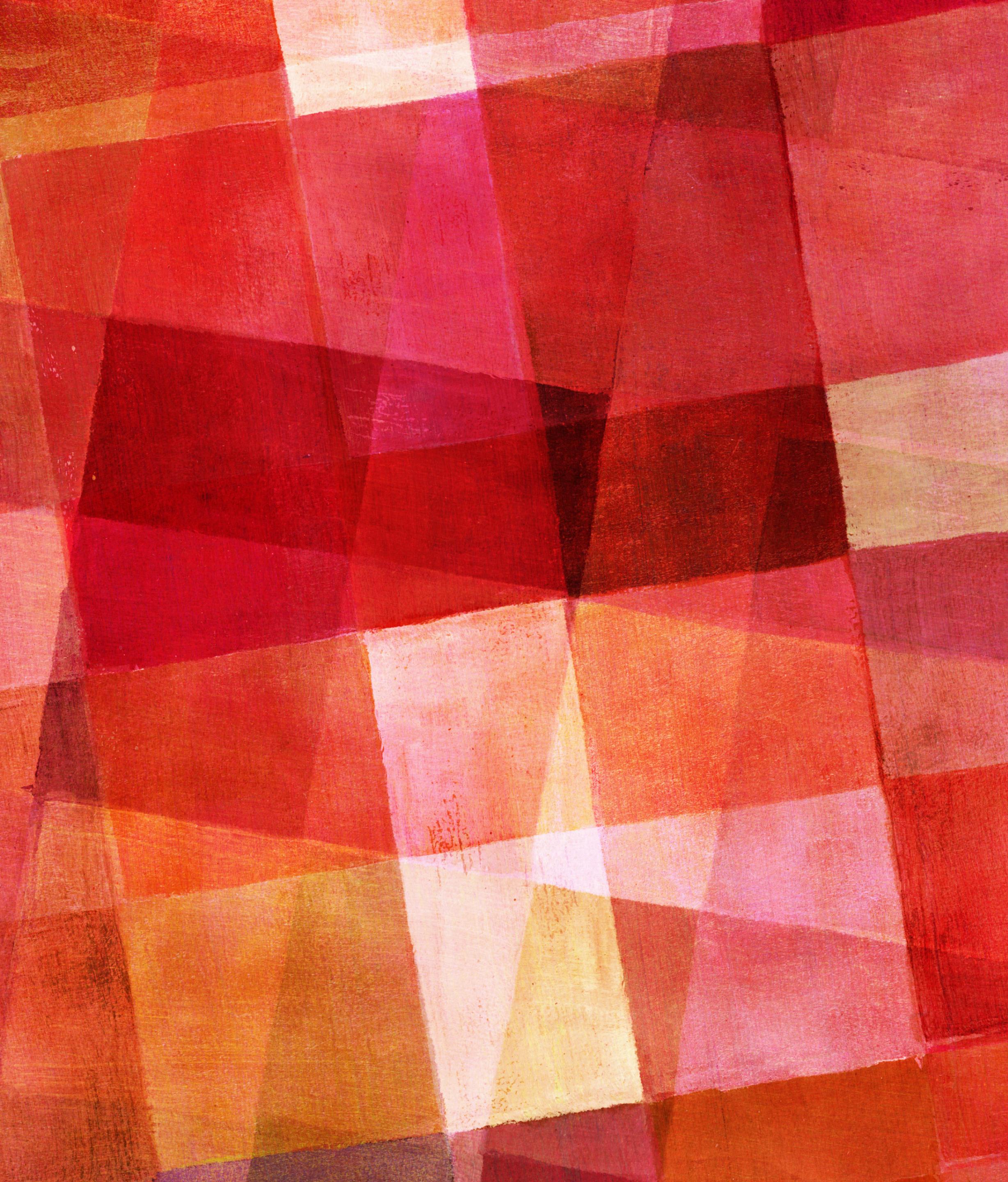
- Profile Picture:** Focusing
- Name:** Steven Roberts
- Email:** sr320
- Edit profile**
- University of Washington:** University of Washington
- Contact:** sr320@uw.edu
- Website:** robertslab.info
- PRO:** PRO
- Organizations:** (Icons for various organizations)

Follow: @RobertsLab, @course-fish274-2019, @urol-e5, More



REMINDER

Start the Zoom!

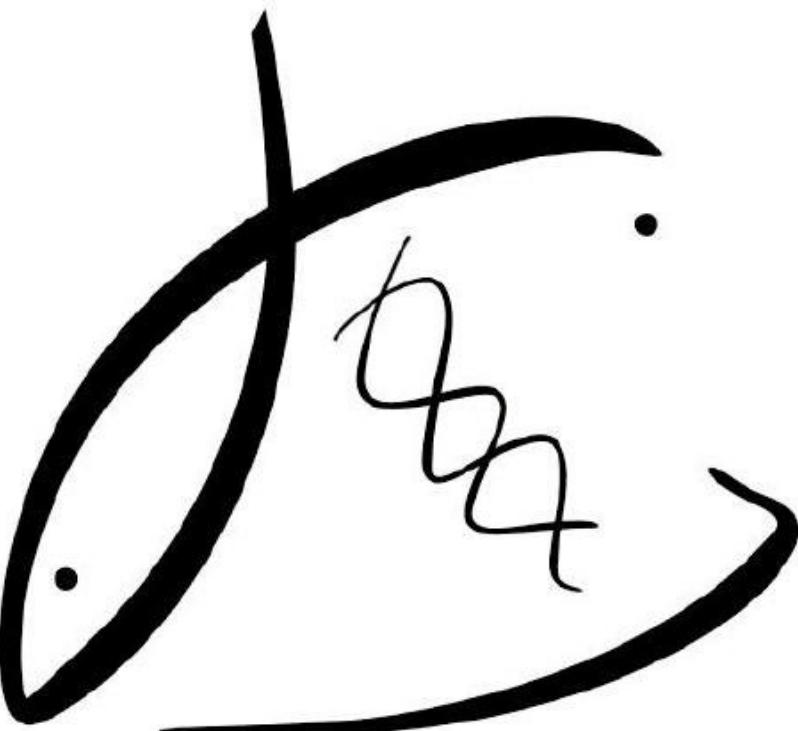


TODAY

- Open Science, Open Data, Reproducibility.
- Try to touch on the entire research cycle and maybe beyond (teaching, administration, service...)
- By no means perfect, and there are plenty of examples of “better”, this is just what has *evolved* to work for me.

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
[@sr320](https://twitter.com/sr320)

Open Access Week

- You are free to Share!

These slides plus links etc @
github.com/sr320/talk-UW-OA-2016

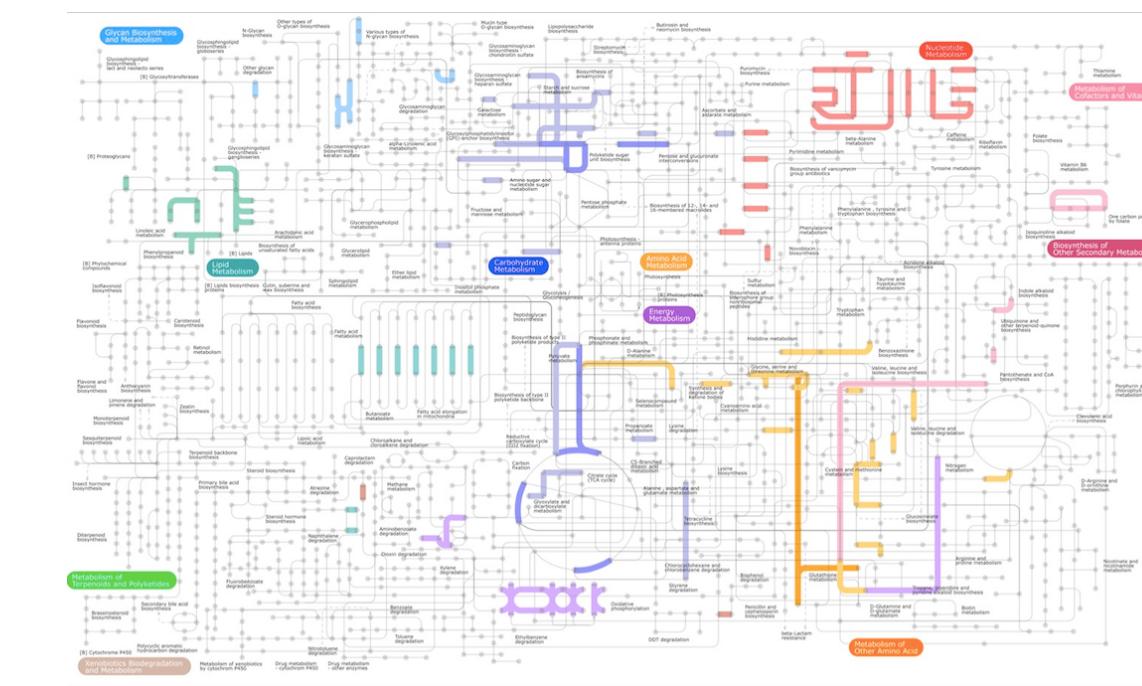
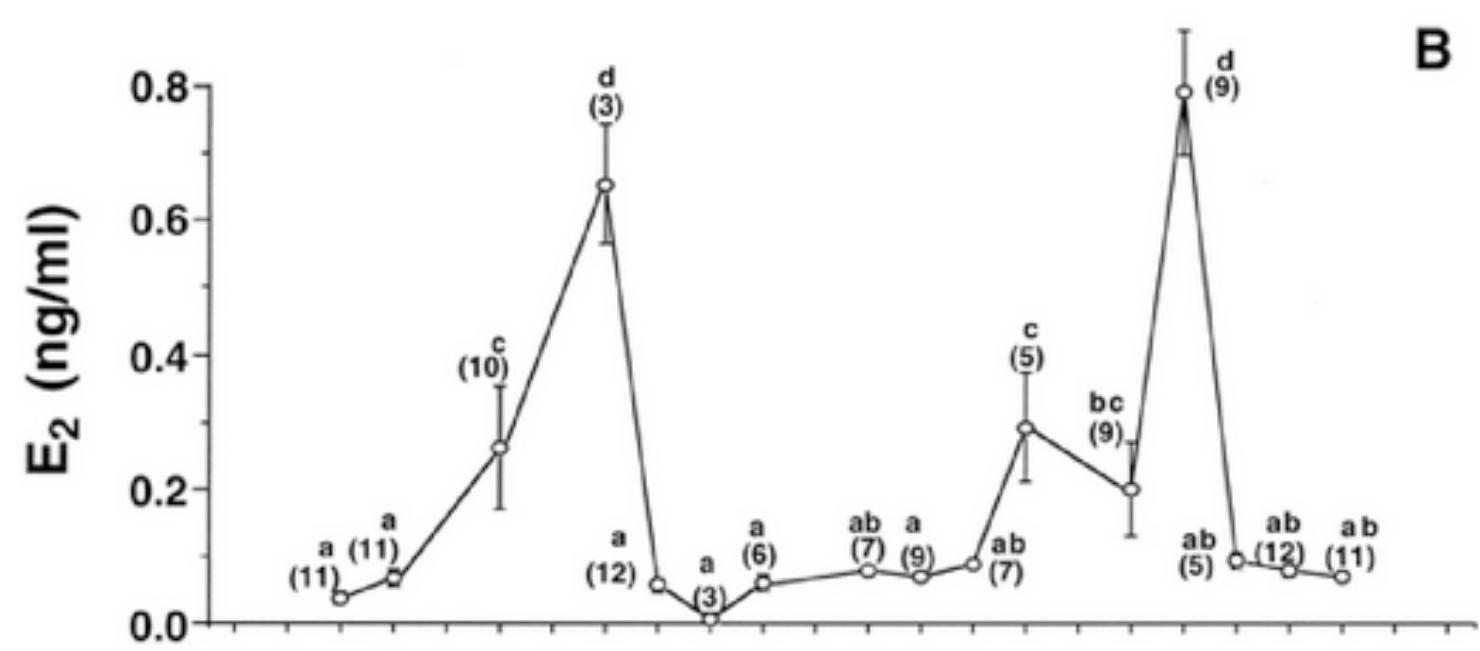
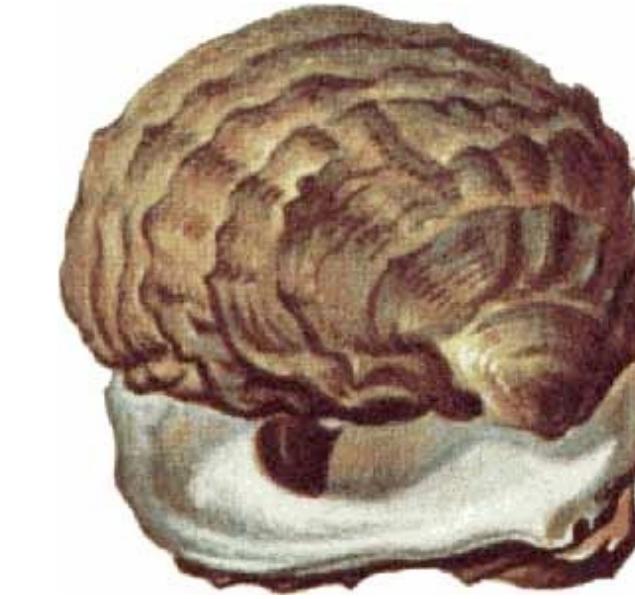
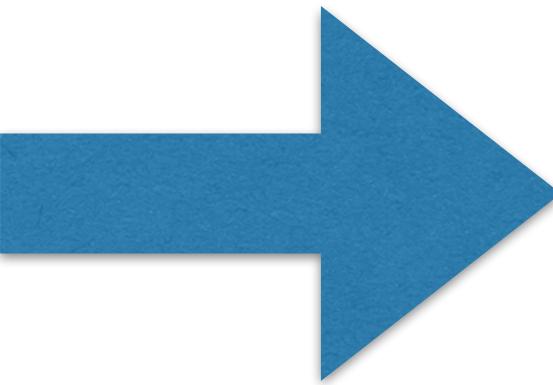
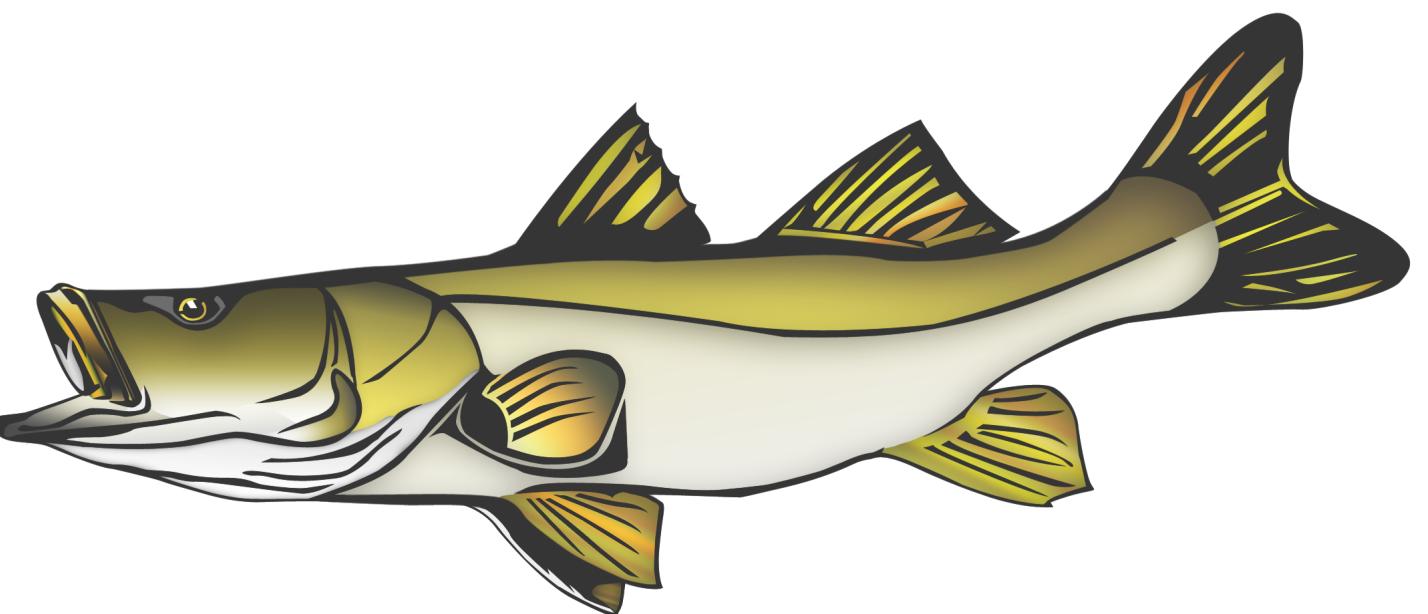
Outline

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

Context

Physiology

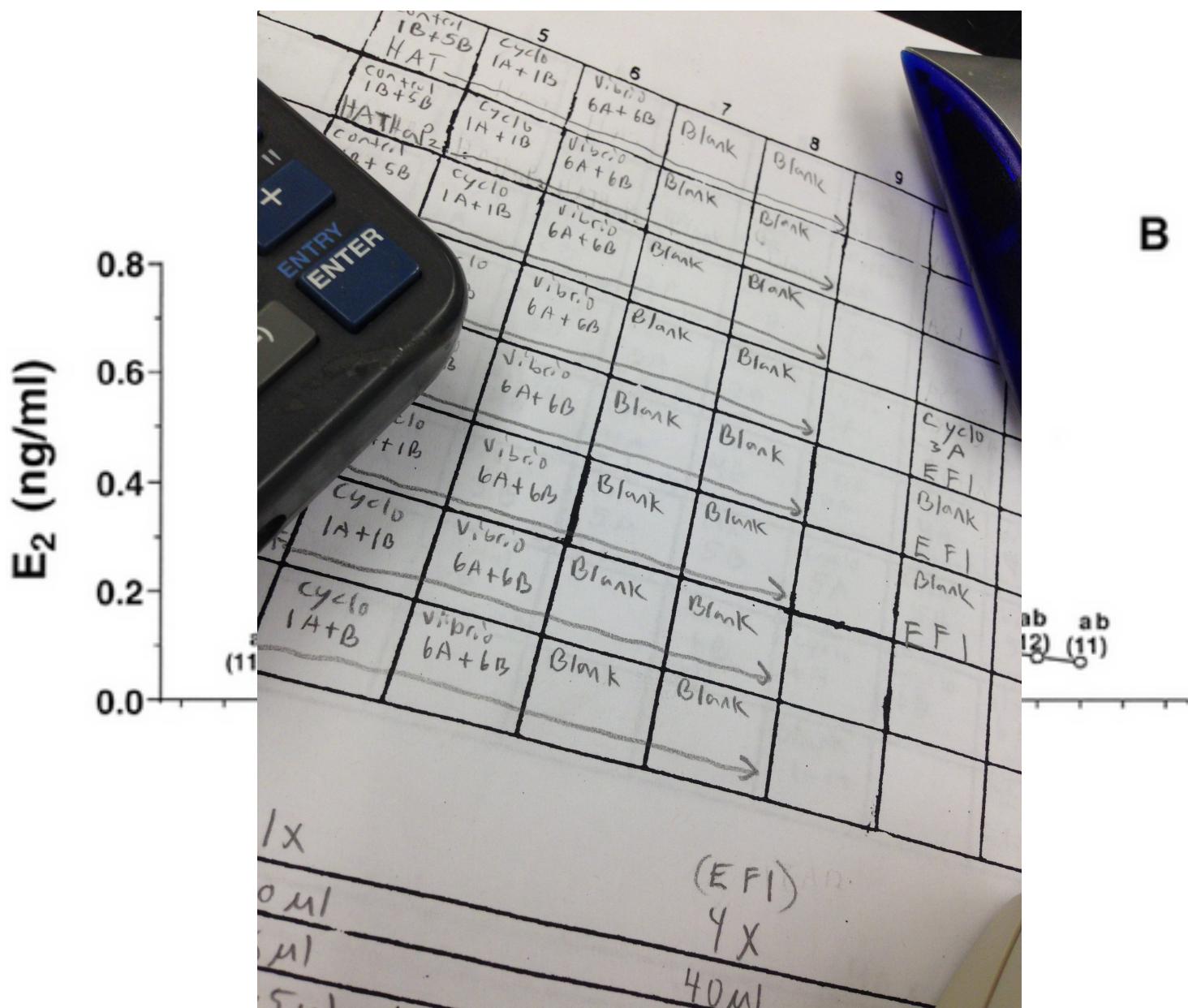
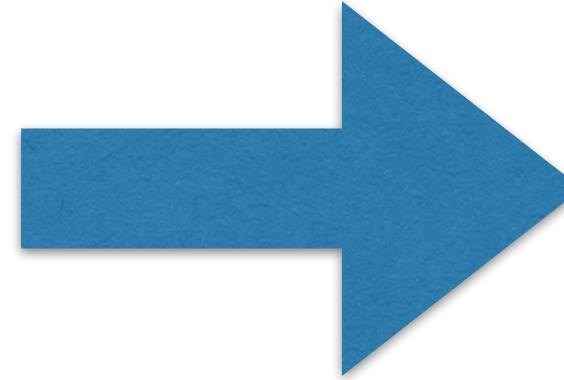
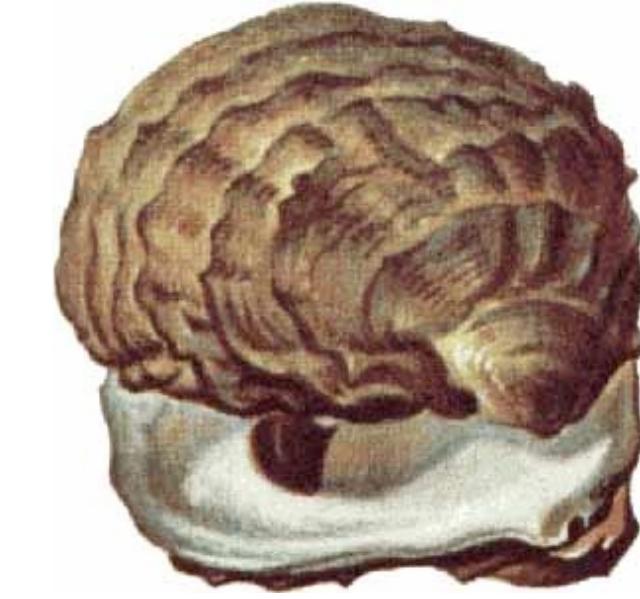
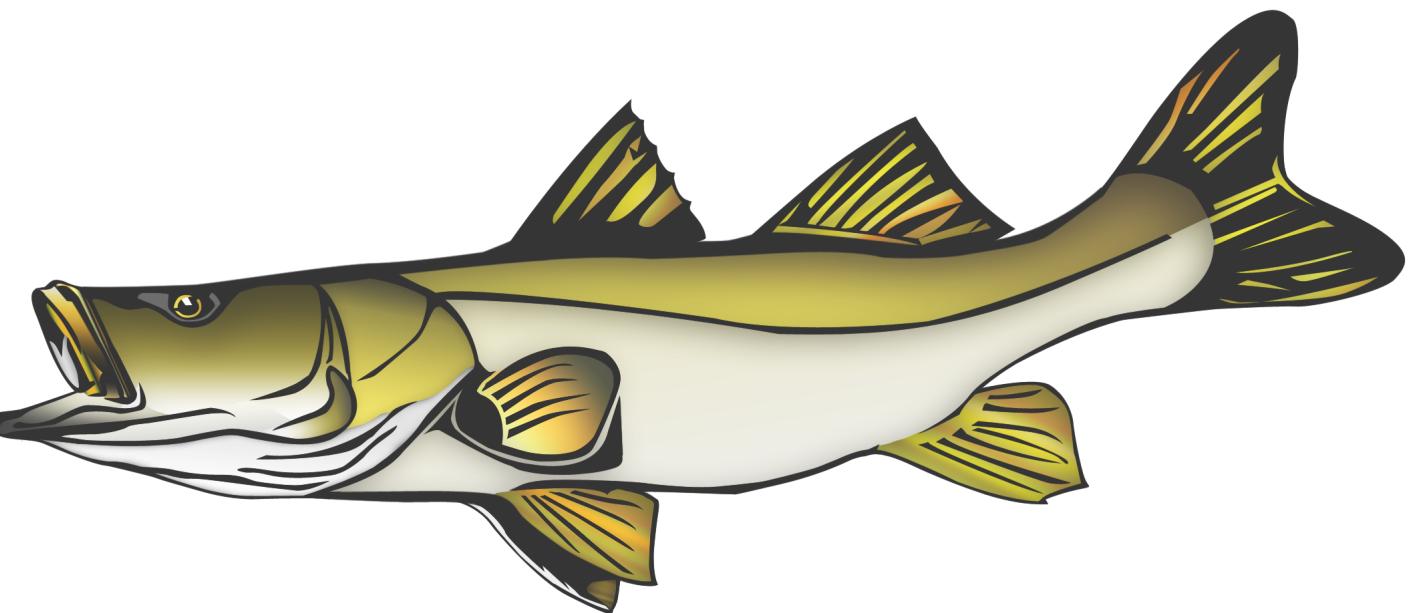
How fundamental processes work in aquatic species



Context

Physiology

How fundamental processes work in aquatic species

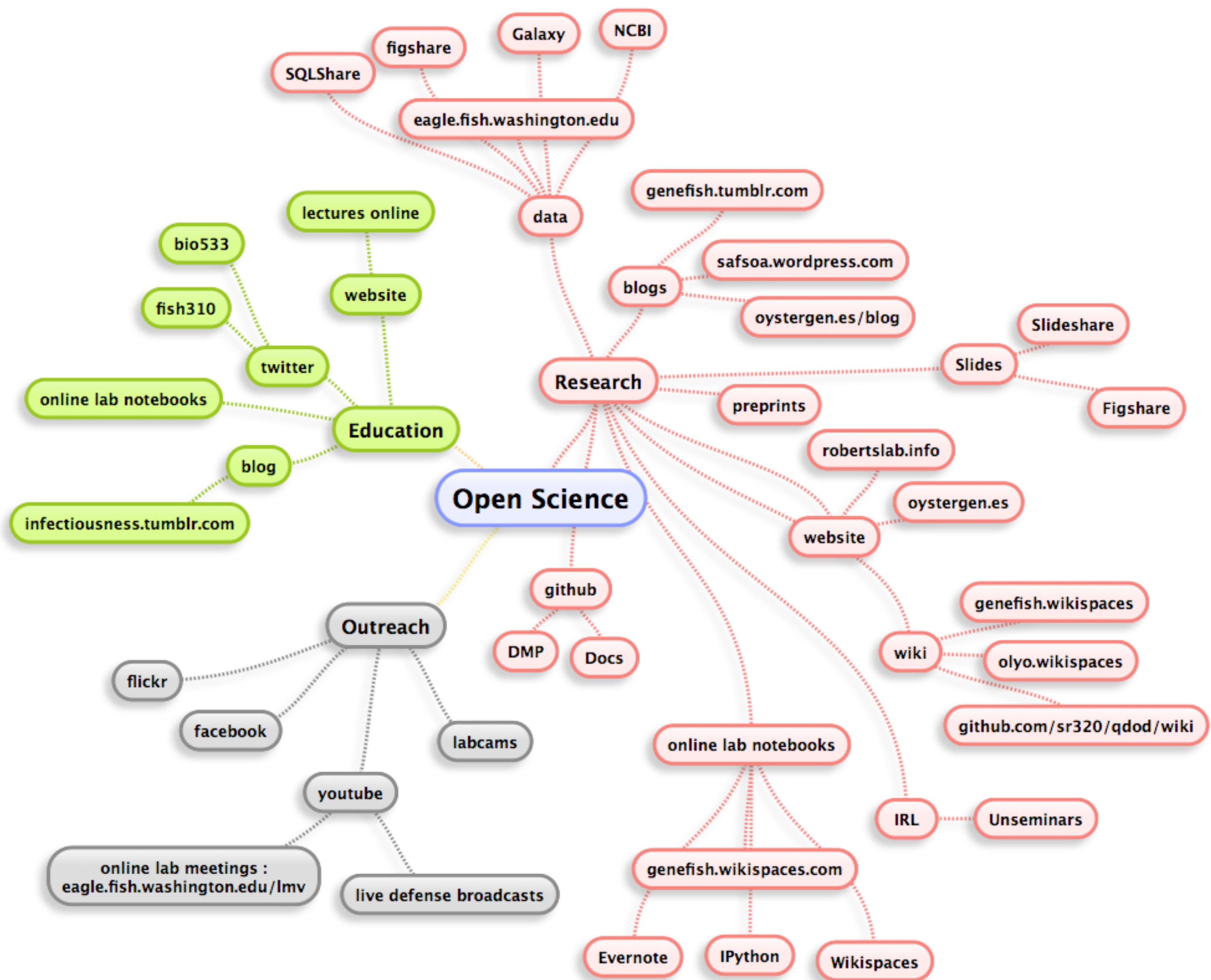


Data Rich

Data is
Open &
Online

*Data is
Open &
Online*

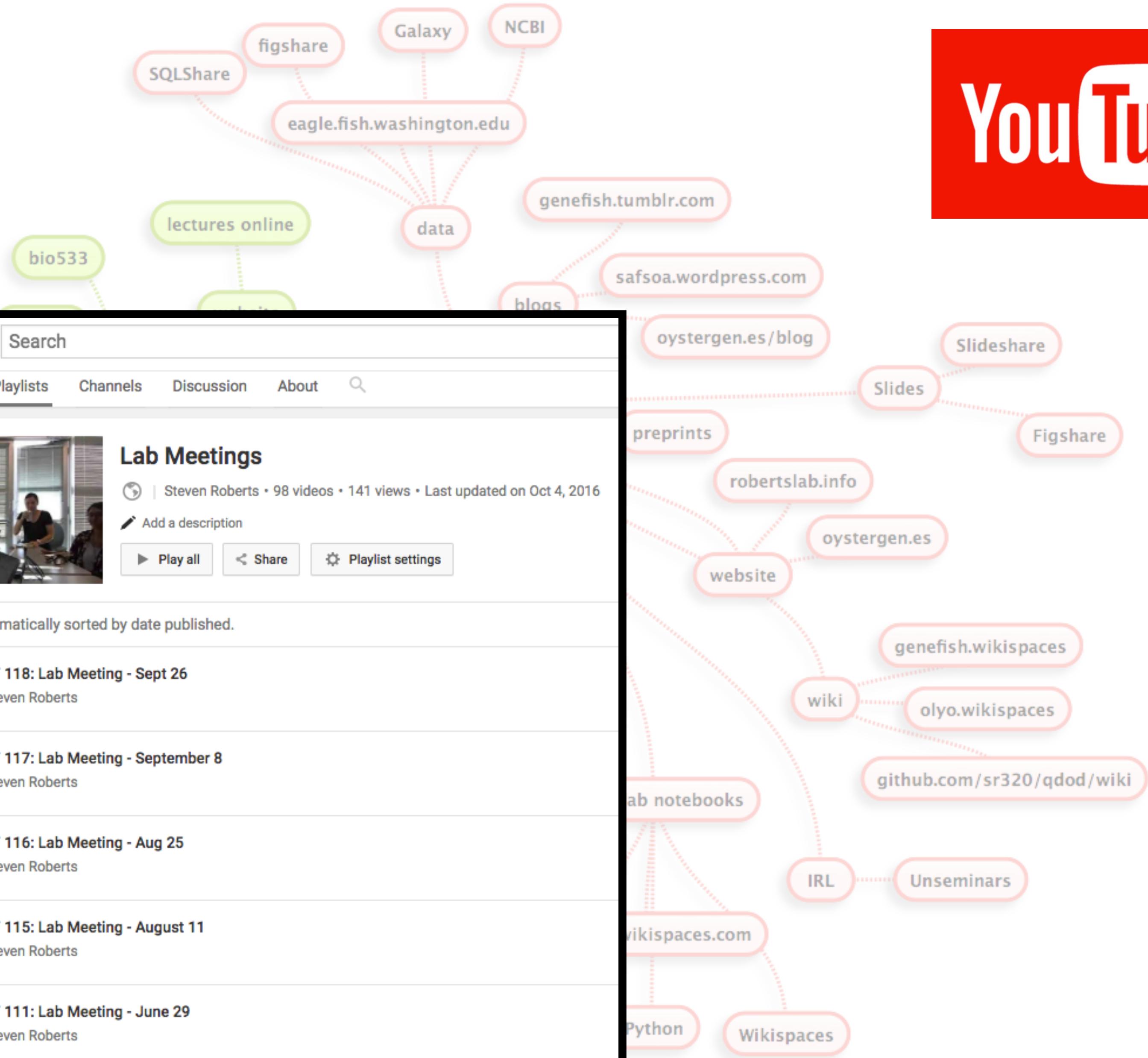
So is everything else



The screenshot shows a YouTube interface with a dark theme. At the top, there's a navigation bar with 'Home', 'Videos', 'Playlists', 'Channels', 'Discussion', 'About', and a search bar. Below the navigation is a thumbnail for a video titled 'TWIF 118: Lab Meeting - Sept 26' by Steven Roberts. The main content area is titled 'Lab Meetings' and displays a list of six videos:

- 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 WATCHED TWIF 111: Lab Meeting - June 29 by Steven Roberts
- 6 TWIF 110: Lab Meeting - June 8 by Steven Roberts

Below the list, it says 'This playlist is automatically sorted by date published.'



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/

onsnetwork.org

Joined October 2013

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

Pinned Tweet

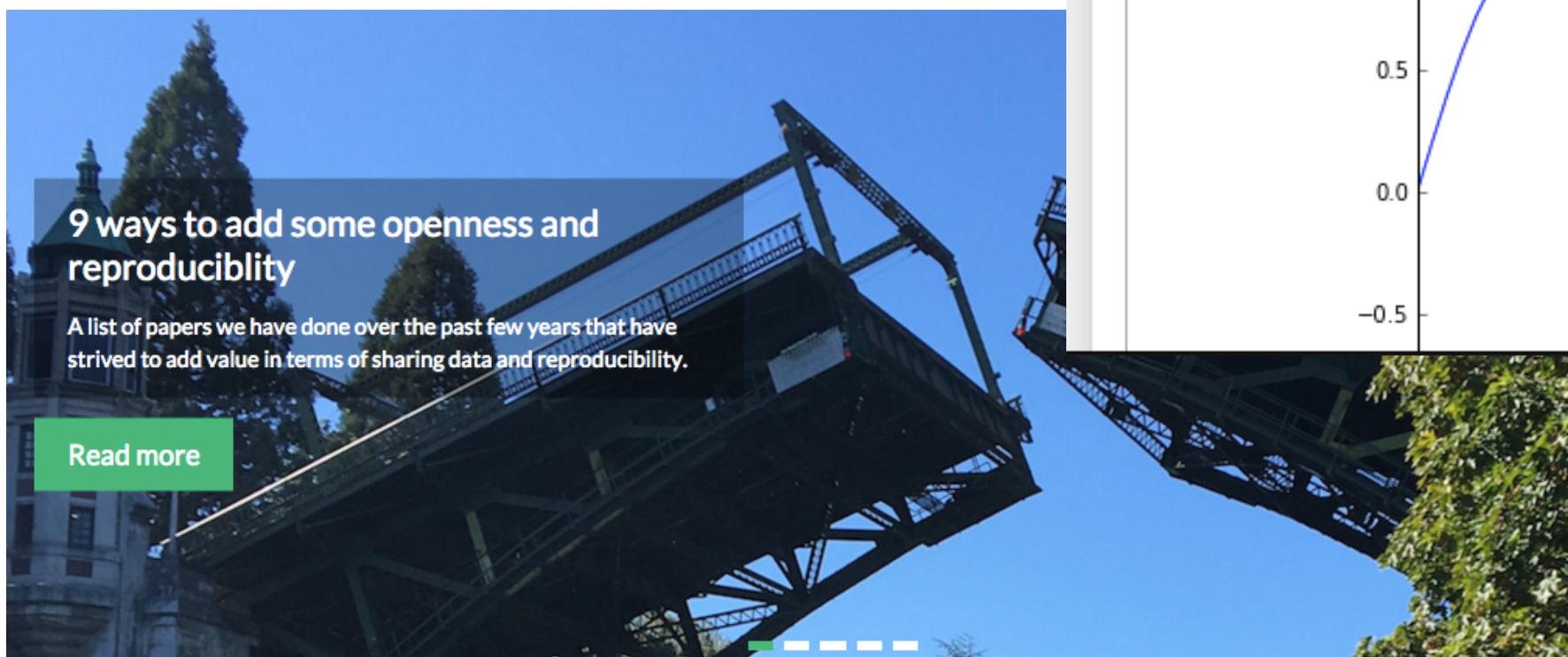
OpenNotebookScience @ONScience · 3 Aug 2014

Really impressed with the open notebooks being kept by @sr320's @EIMD_fhl students onsnetwork.org/eimd/ #openscience as a matter of course

15 9 ...

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

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

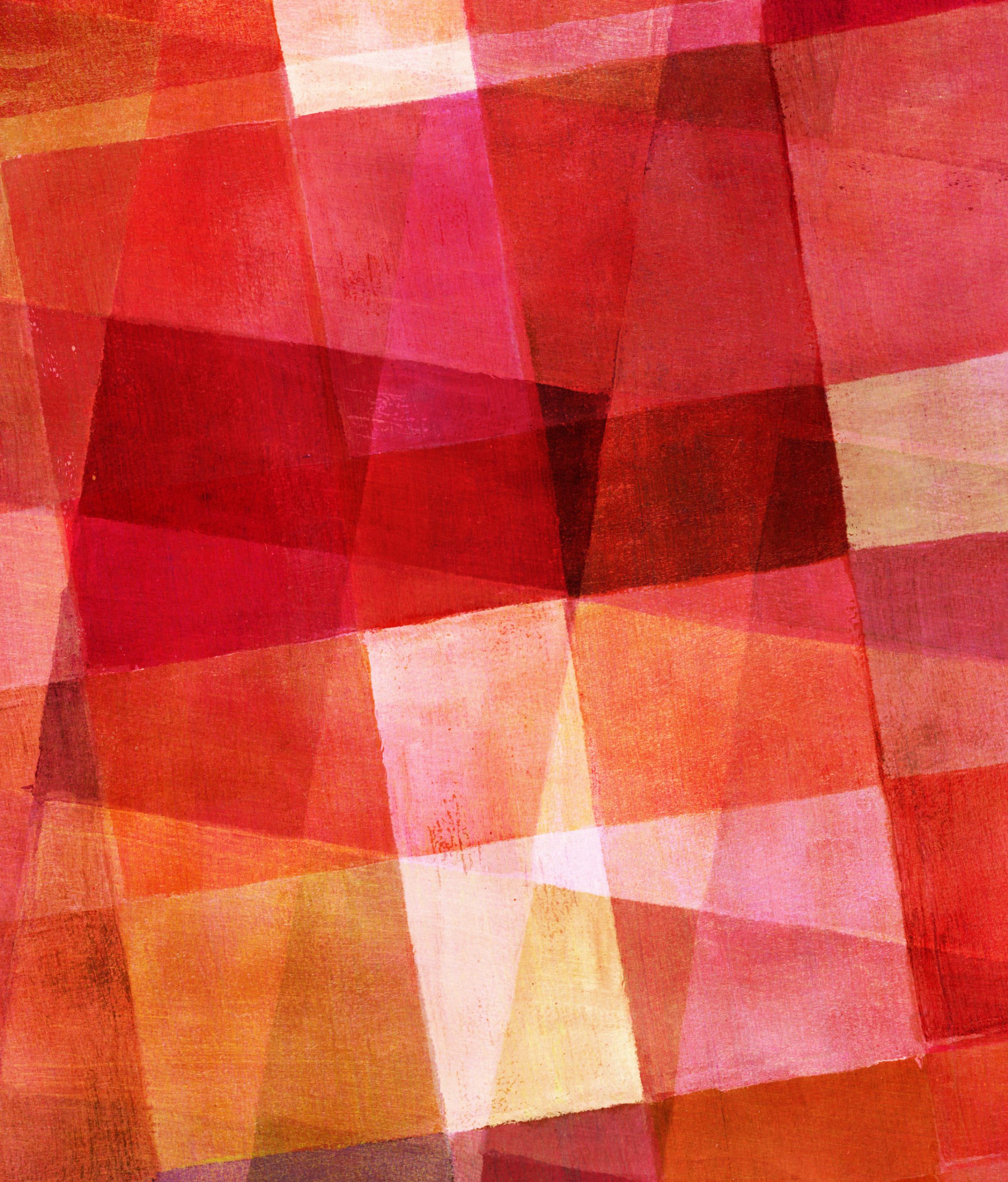
data typo 7 days ago

ipynb Update READMEs 5 days ago

README.md typo 7 days ago

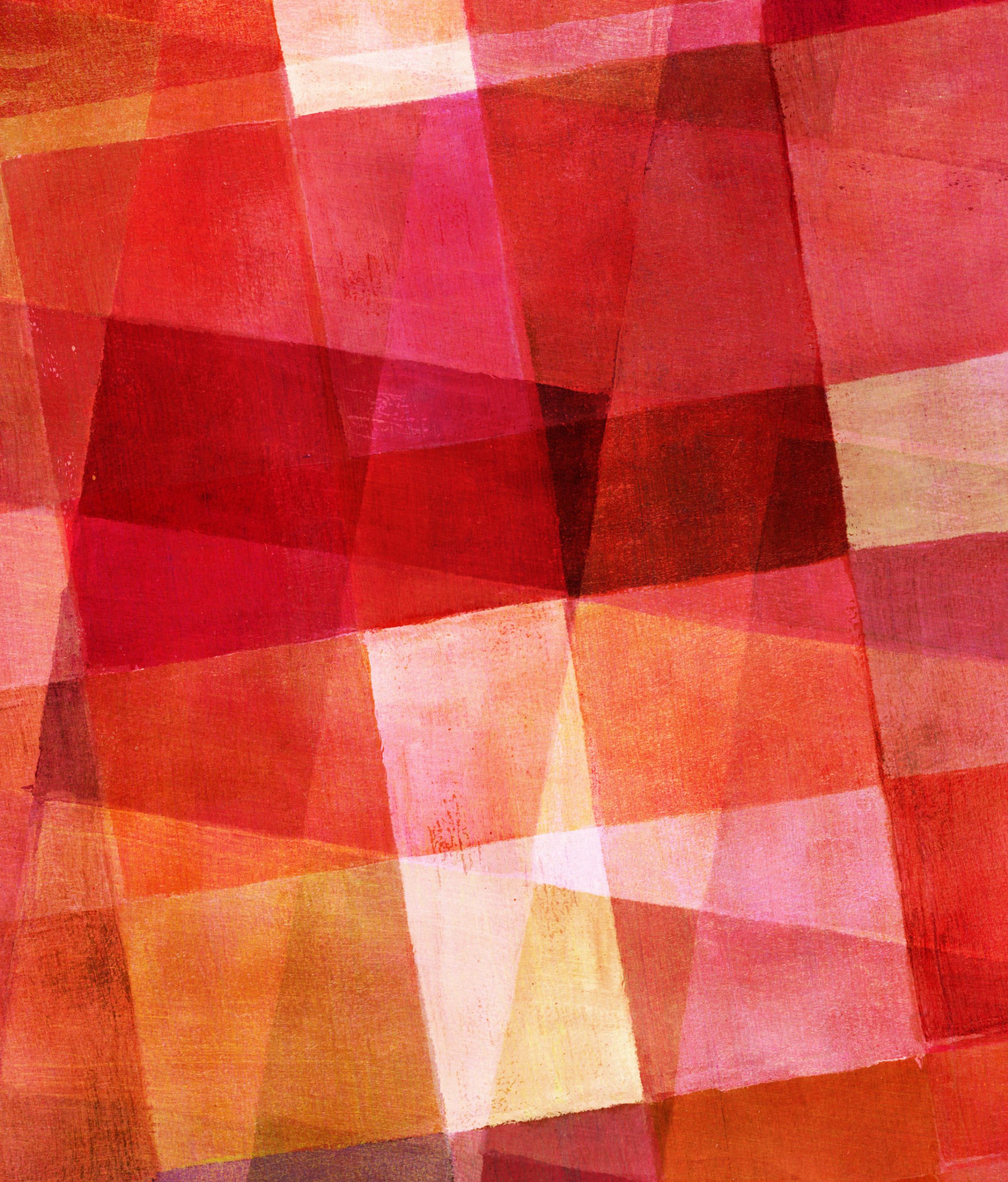
Analyses and data associated with manuscript:
Germline DNA methylation in six species of reef corals: patterns and potential roles in response to environmental change

A GitHub repository page for 'Coral-CpG' showing 22 commits, 1 branch, 0 releases, and 2 contributors. The repository is associated with a paper titled 'Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change'. The repository page includes a list of files and their commit history, and a summary of analyses and data associated with the manuscript.



THE LAB NOTEBOOKS

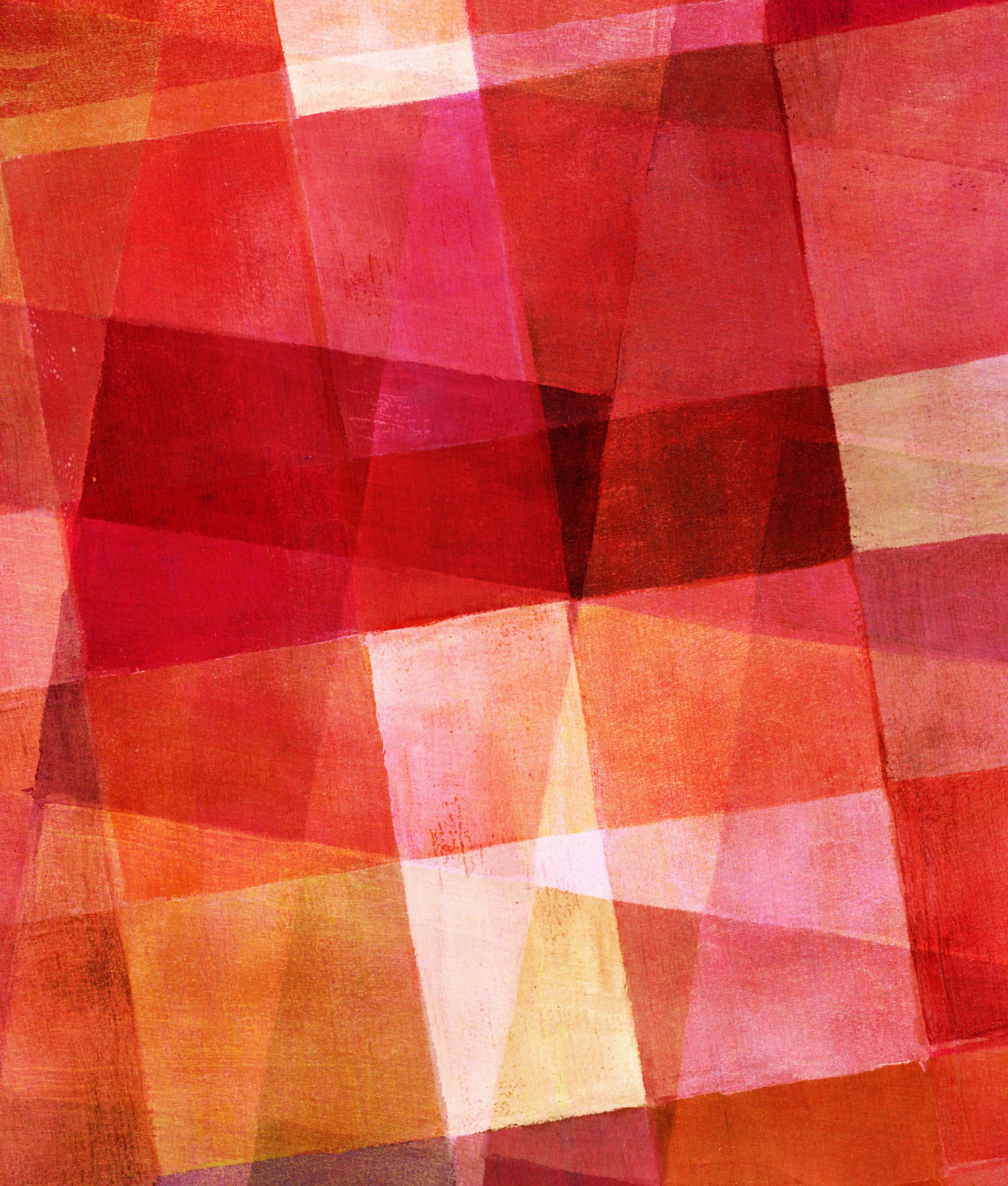
- Probably seen the greatest evolution over the past decade.
- GitHub is at the core, with NAS.
 - Daily crunching (bash, Jupyter, R).
 - Notebook posts.
- Wordpress as a quick an easy solutions and fine aggregator.



THE LAB NOTEBOOKS

- Probably seen the greatest evolution over the past decade.
- GitHub is at the core, with NAS.
 - Daily crunching (bash, Jupyter, R).
 - Notebook posts.
- Wordpress as a quick an easy solutions and fine aggregator.

Live Show and Tell



DID I HIT EVERYTHING?

- Jupyter
- NAS
- Wordpress
- Atom
- GitHub
 - nb-2020
 - “notebook” - Markdown
- RStudio - Git Integration
- Pro tips
- Droplr
- Monosnap

The background of the left half of the image features a textured, abstract design composed of overlapping circles in various shades of green, blue, and yellow. The circles are semi-transparent, creating a layered effect against a dark, mottled background.

MORE WITH GITHUB

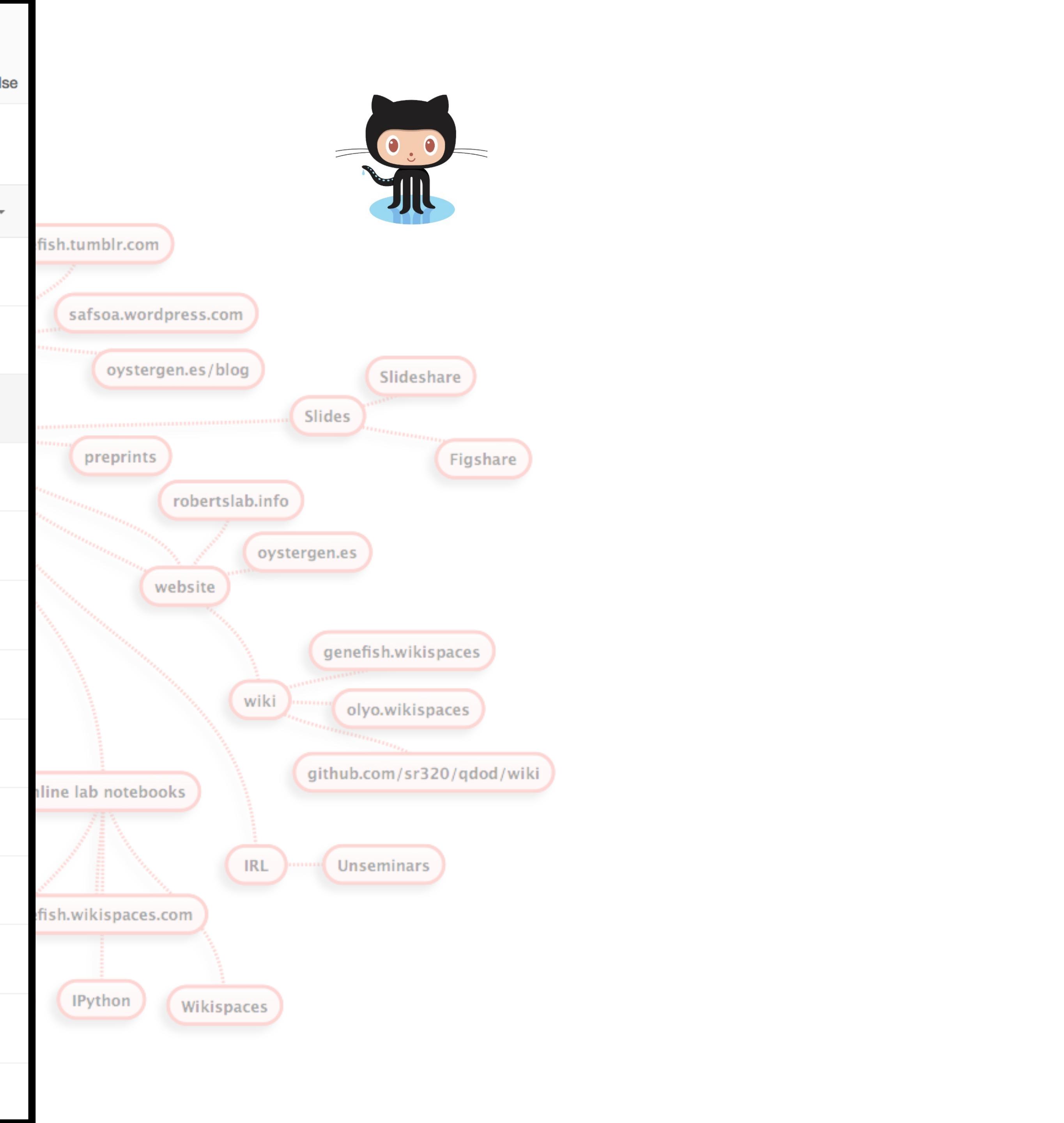
sr320 / LabDocs

Code Issues 20 Pull requests 0 Projects 1 Wiki Pulse

Filters is:issue is:open Labels Milestones

20 Open ✓ 280 Closed Author ▾

- ① **Develop Schematic of Entirety of Oly Oyster Samples**
#317 opened an hour ago by sr320
- ① **Proteomics Meeting - Developing Timeline**
#314 opened a day ago by sr320
- ① **Hyak node**
#313 opened a day ago by sr320
- ① **Take action on genomes..**
#310 opened 2 days ago by sr320
- ① **System update available for Synology servers.** enhancement
#308 opened 2 days ago by kubu4
- ① **R Studio Server broken** Docker enhancement low priority
#307 opened 2 days ago by kubu4
- ① **Do we modify all xserves for use?** labmeeting
#299 opened 8 days ago by sr320
- ① **Plan to handle extended power outage?**
#295 opened 14 days ago by kubu4
- ① **Why does a post not break to read more with Jekyll-now**
#292 opened 15 days ago by sr320
- ① **Consider how to best get tag data online**
#288 opened 18 days ago by sr320
- ① **Proposal for Data Incubator?** labmeeting
#285 opened 22 days ago by sr320
- ① **Deploy Shellfish out to BBC** labmeeting
#277 opened on Sep 26 by sr320
- ① **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](#)

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	Questions
one	Bash, version control, Project Set-up	Setting Up and Managing a Bioinformatics Project 21-35; Remedial Unix Shell 37-54	Questions
two	Jupyter, Annotation	Retrieving Bioinformatics Data 109-124, Unix Data tools 125-168	Questions
three	Projects	Working with Sequence Data 339-354	Questions

[Edit](#) [New Page](#)
[Pages 3](#)
[Week 3](#)

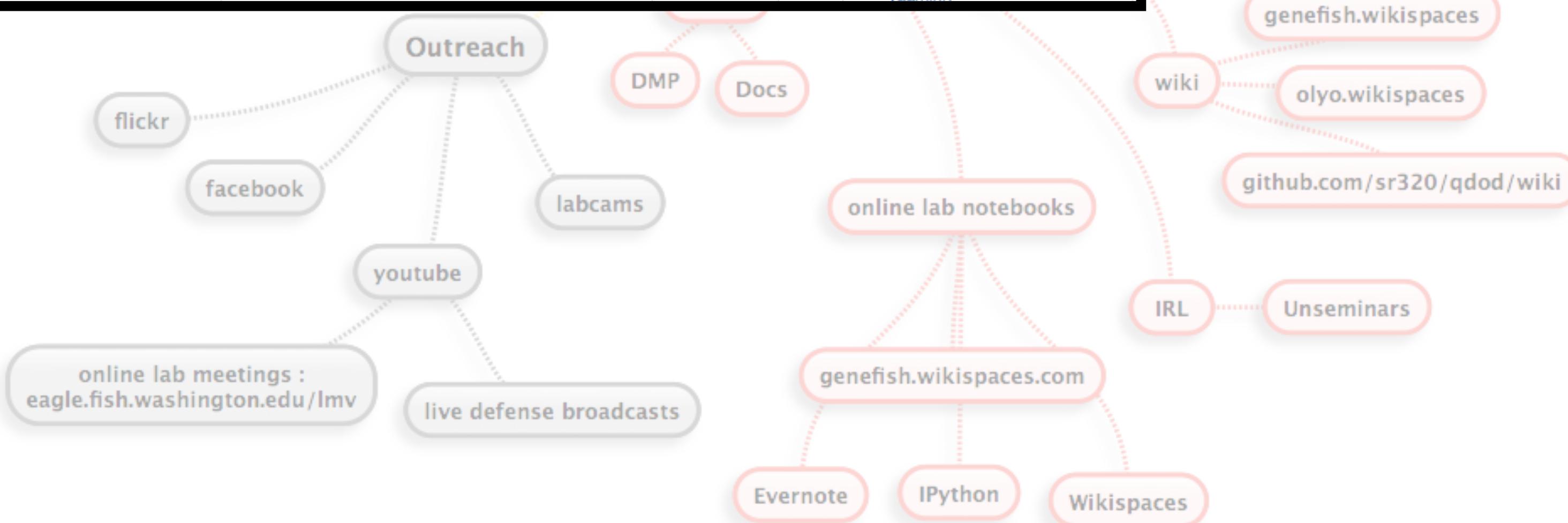
- Notes
- Quiz
- Project

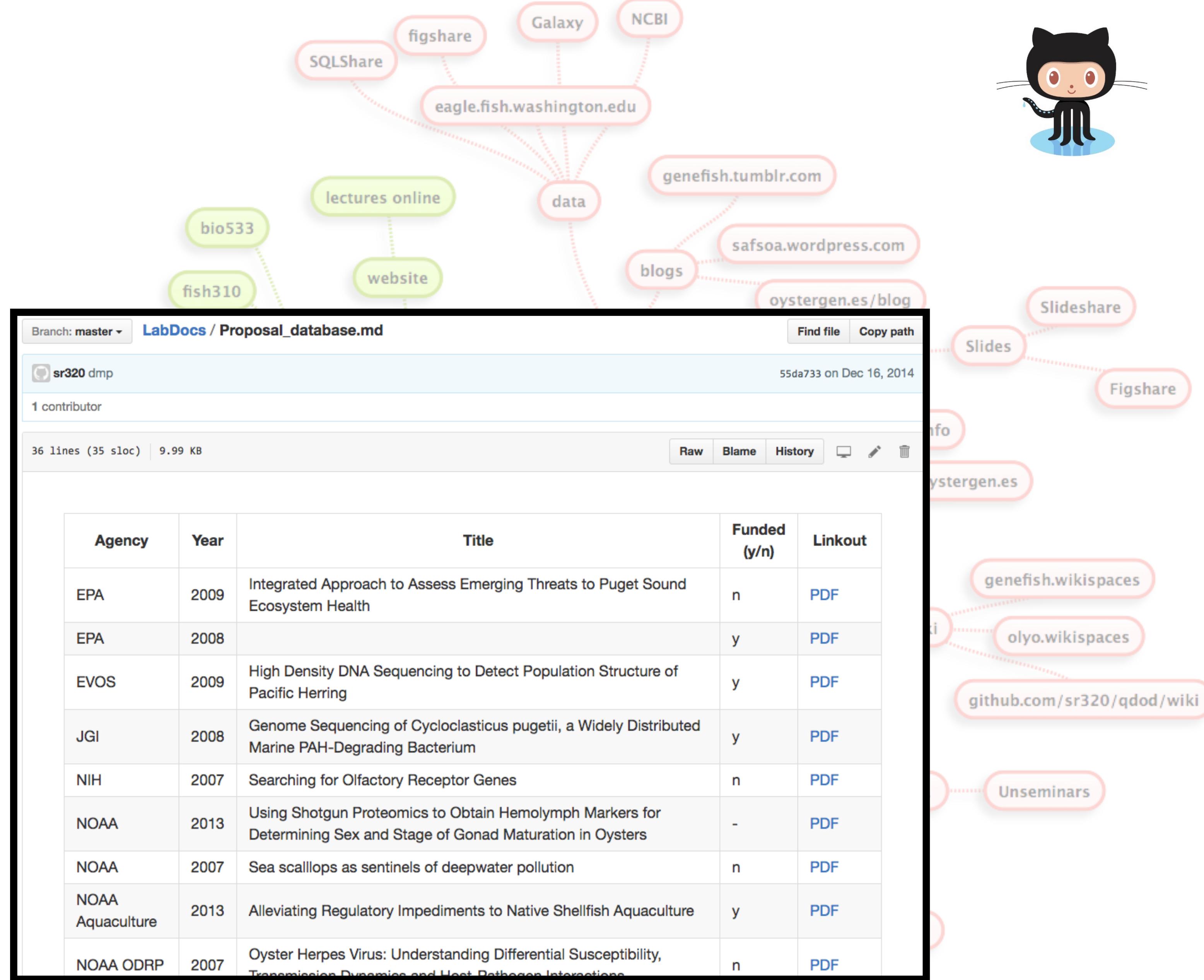
[Week 4](#)

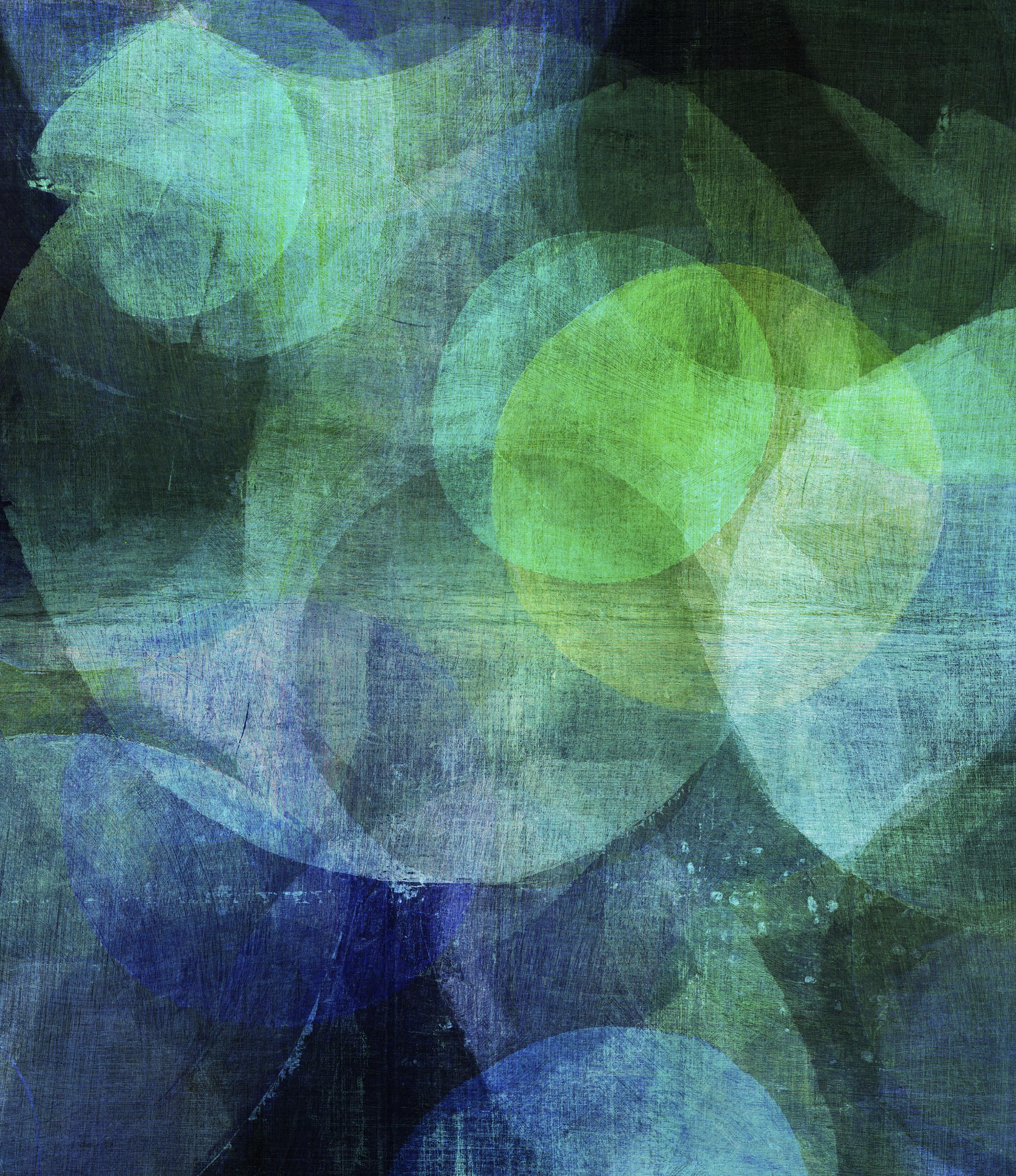
- Notes
- Quiz
- Project

[Grades](#)
[Class Repos](#)

- hputnam
- vaaminiy







MORE WITH GITHUB

Live Show and Tell



DID I HIT EVERYTHING?

- Personal
- Organizations (Lab, OARS)
- Teaching
- Issues!
- Wiki
- RobertsLab : Resources



MANUSCRIPTS & OTHER PRODUCTS



MANUSCRIPTS & OTHER PRODUCTS

Live Show and Tell



MANUSCRIPTS & OTHER PRODUCTS

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.

“

<https://github.com/sr320/talk-UCSB-repro>