FREE SOFTWARE FOR ENVIRONMENTAL SCIENCES

Weiming Hu

Center for Western Weather and Water Extremes
Scripps Institution of Oceanography
UC San Diego

March 19, 2023

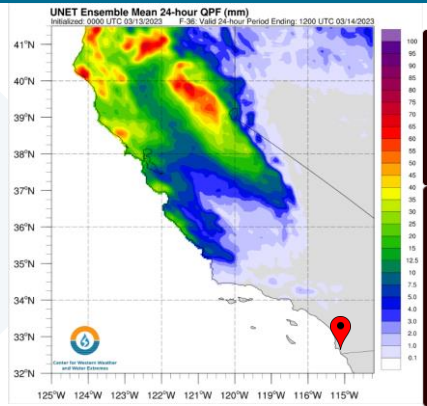
R/V *Sally Ride*, Anacortes , Washington. *Credit: Jeff Dillon.*

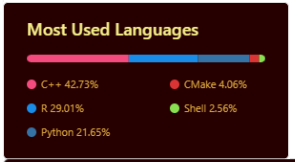
BACKGROUND

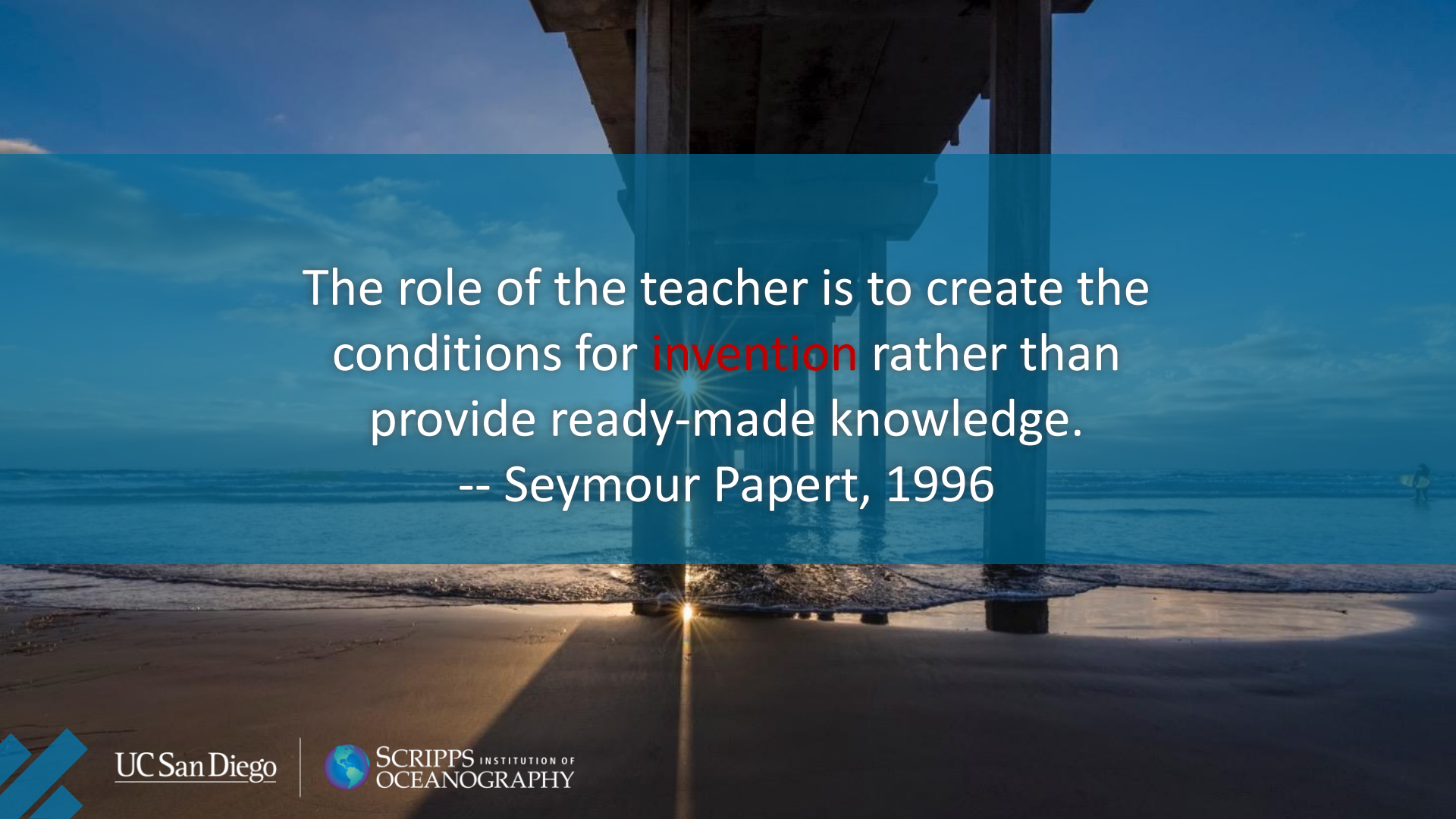
- Educator
- Geographer
- Machine Learning Scientist

- Open-Source Developer
- Free Software Enthusiast

- Electric Unicycle (EUC) Rider
- Guitar







UC San Diego | SCRIPPS INSTITUTION OF OCEANOGRAPHY

OUTLINE

- Motivation
- Open Science
- Open Access
- Free Software for Open Science
- Teaching Free Software for Environmental Science
- What Can the Dev Community do

MOTIVATION

1. Teaching commercial software in public institutions
2. Using commercial software in public institutions

MOTIVATION

This talk is not about:

- Specific software
- Science topic
- Pointing fingers

This talk is about:

- Similarities between science and free software
- Understanding people’s decisions
- Change and better practice

OPEN SCIENCE

Definition of *Open* in Open Data, Open Content, and Open Knowledge:

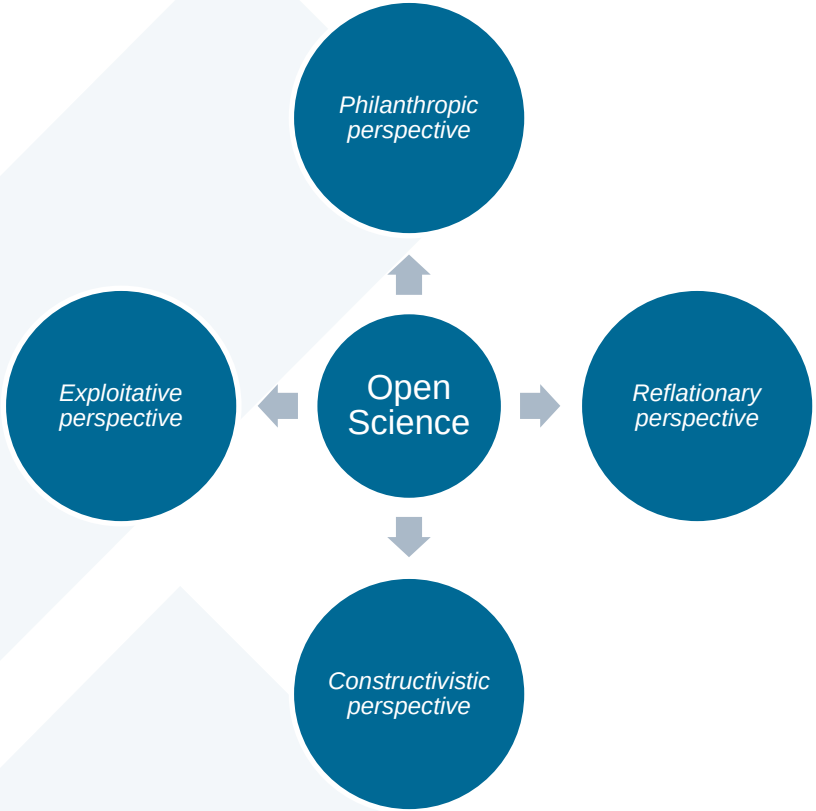
> “Open means anyone can freely access, use, modify, and share for any purpose.” [1]

Definition of *Free* in Free Software:

> “Free software means software that respects users' freedom and community. Roughly, it means that the users have the freedom to run, copy, distribute, study, change and improve the software.” [2]

[1] Open Knowledge Foundation, https://opendefinition.org/
[2] GNU Operating System, https://www.gnu.org/philosophy/free-sw.en.html

PERSPECTIVES OF OPEN SCIENCE

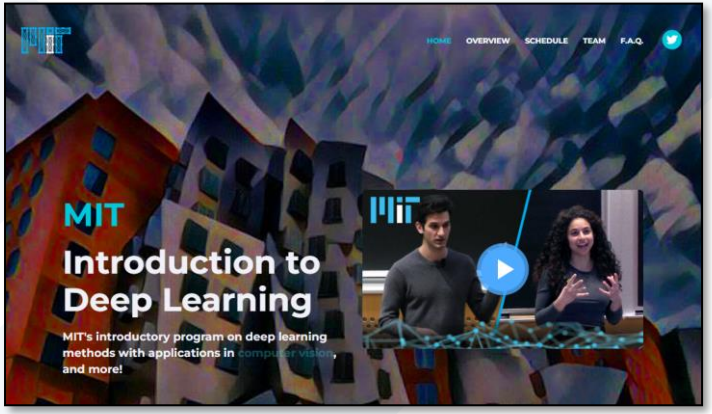


Friesike, S., Widenmayer, B., Gassmann, O. and Schildhauer, T., 2015. Opening science: towards an agenda of open science in academia and industry. *The journal of technology transfer*, *40*, pp.581-601.

PERSPECTIVES OF OPEN SCIENCE

Philanthropic Perspective

- Public lectures or courses

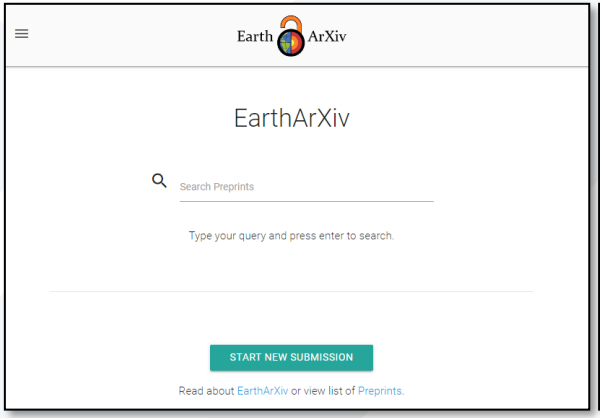


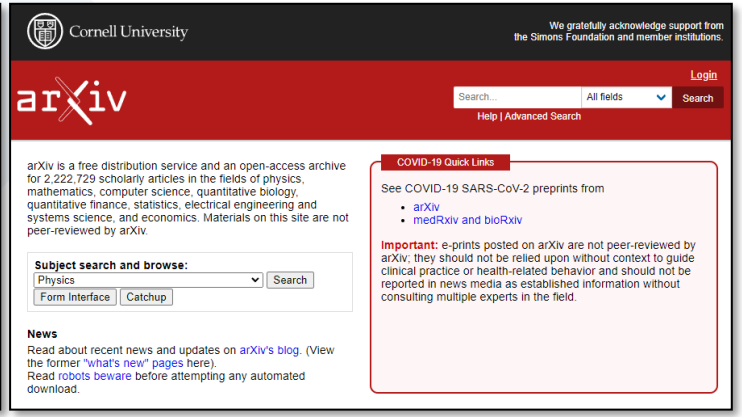
- Open access journals

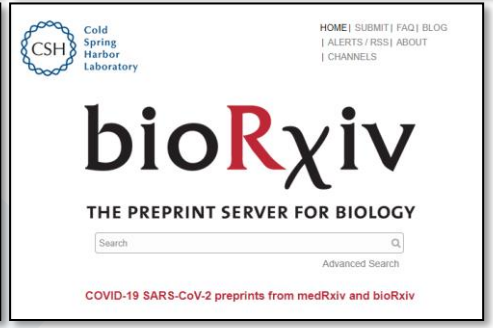
PERSPECTIVES OF OPEN SCIENCE

Reflationary perspective

- making scientific results freely available during pre-publication.







PERSPECTIVES OF OPEN SCIENCE

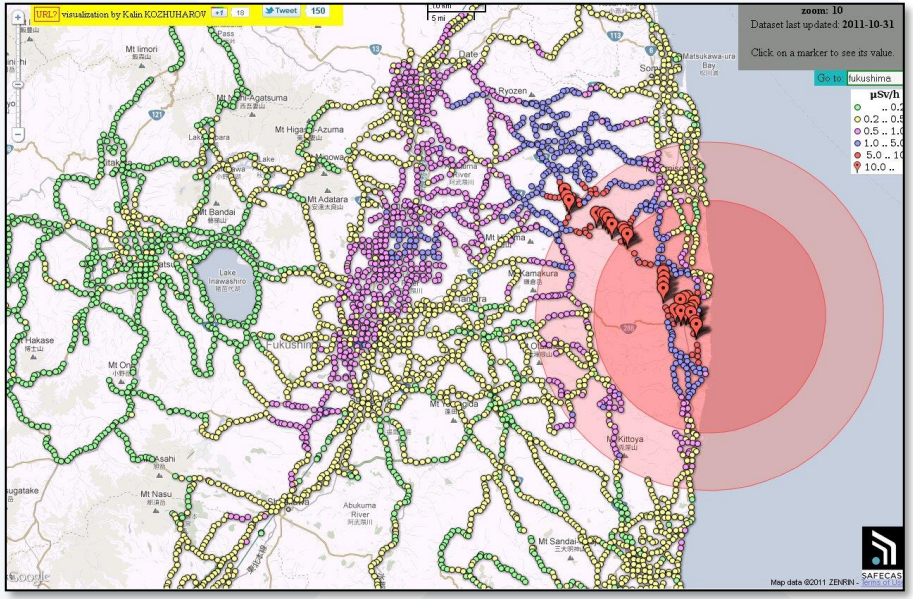
Constructivistic perspective

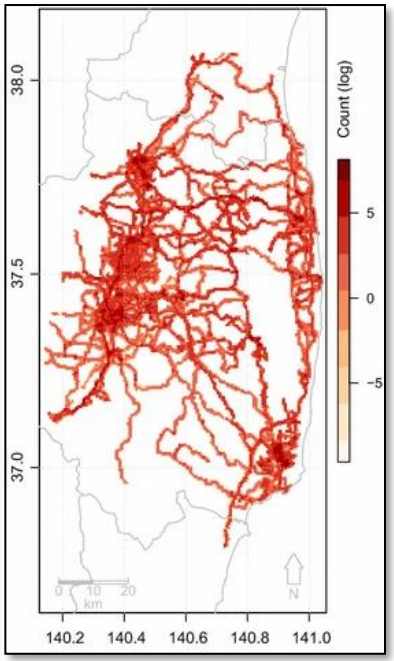
- Crowdsourcing & interdisciplinary

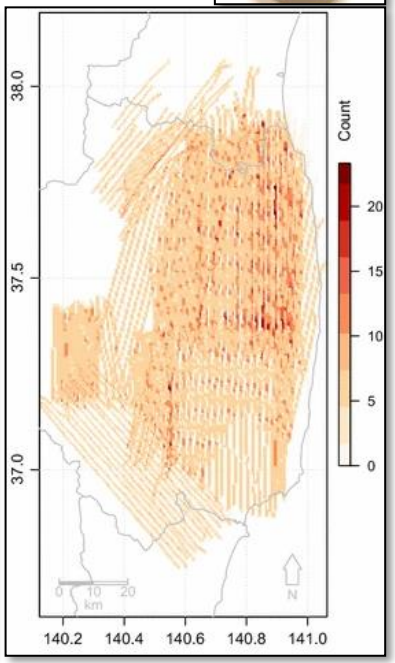
PERSPECTIVES OF OPEN SCIENCE

Constructivistic perspective

- Crowdsourcing & interdisciplinary







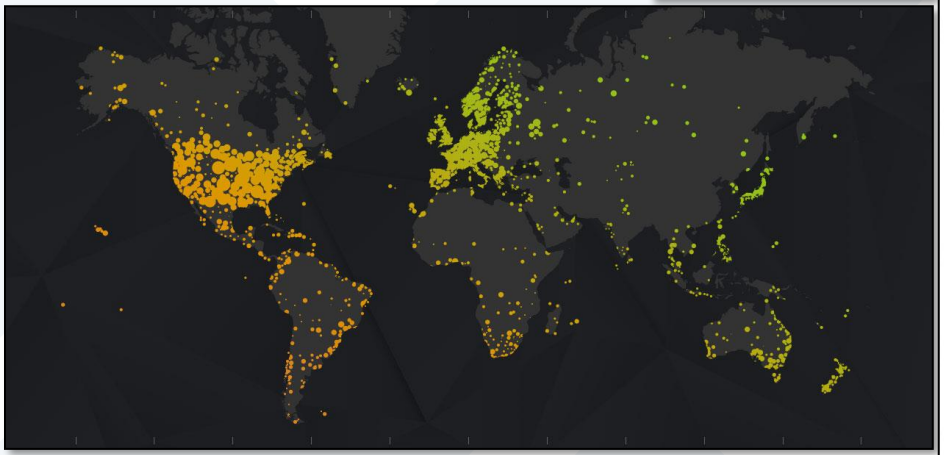
Safecast collects more than a million points of data: https://diy-scib.org/blog/safecast-collects-more-million-points-data

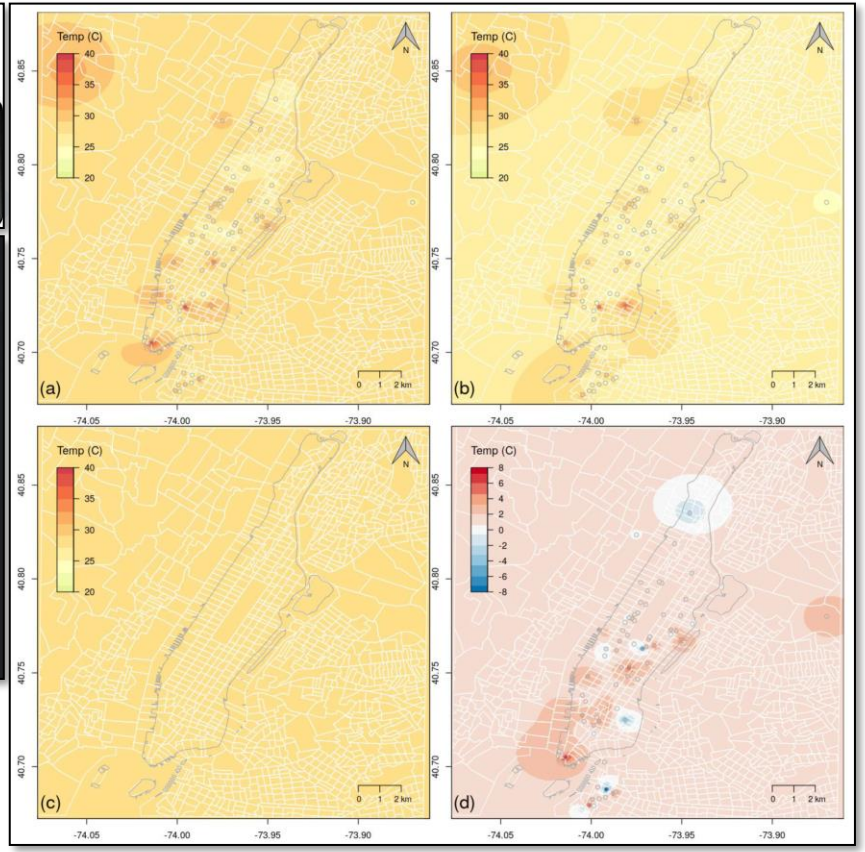
Hultquist, C. and Cervone, G., 2018. Citizen monitoring during hazards: Validation of Fukushima radiation measurements. *GeoJournal*, *83*(2), pp.189-206.

PERSPECTIVES OF OPEN SCIENCE

Constructivistic perspective

- Crowdsourcing & interdisciplinary





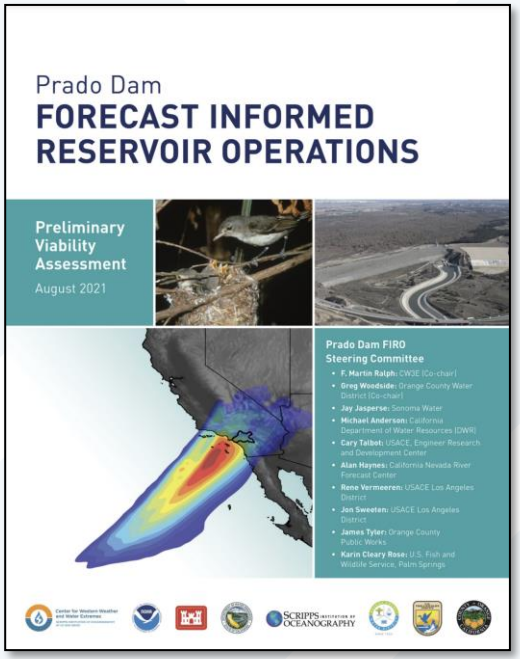
Personal Weather Station Network: https://www.wunderground.com/pws/overview

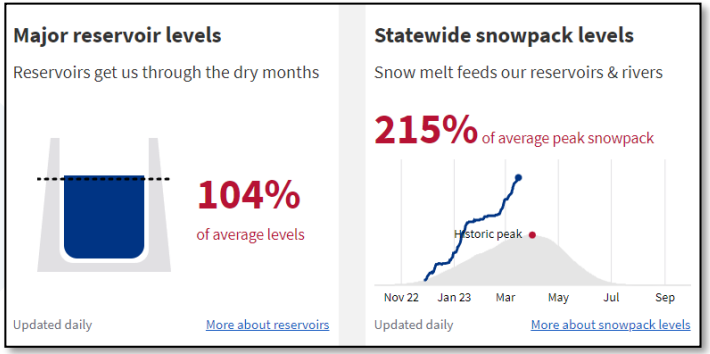
Calovi, M., Hu, W., Cervone, G. and Delle Monache, L., 2021. NAM-NMM Temperature Downscaling Using Personal Weather Stations to Study Urban Heat Hazards. *GeoHazards*, *2*(3), pp.257-276.

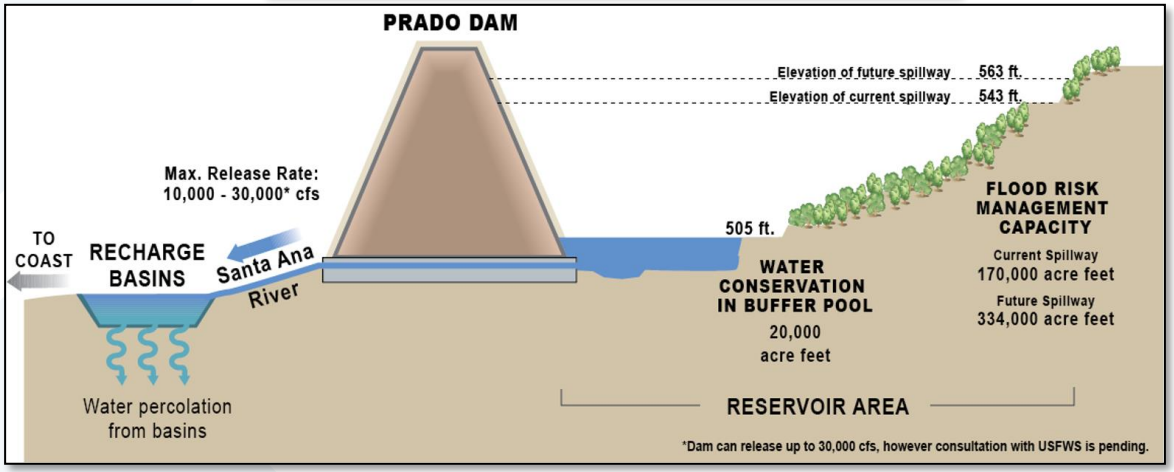
PERSPECTIVES OF OPEN SCIENCE

Exploitative perspective

- application-oriented knowledge



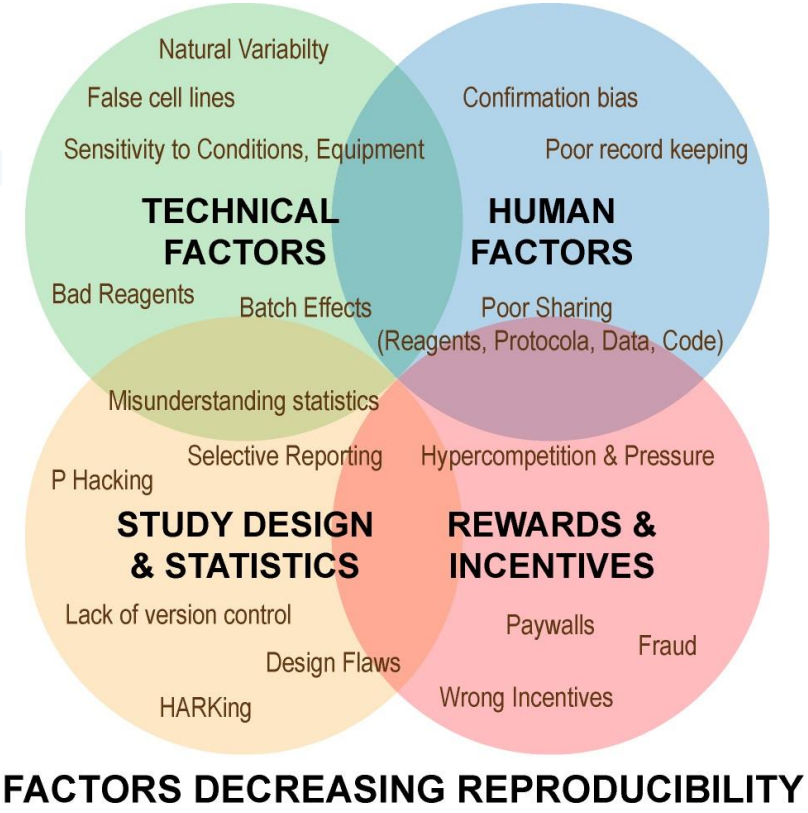




California Water Watch: https://cww.water.ca.gov/
Prado Dam Preliminary Viability Assessment: https://cw3e.ucsd.edu/firo_prado_dam_pva/

WHY DO WE NEED OPENNESS IN SCIENCE

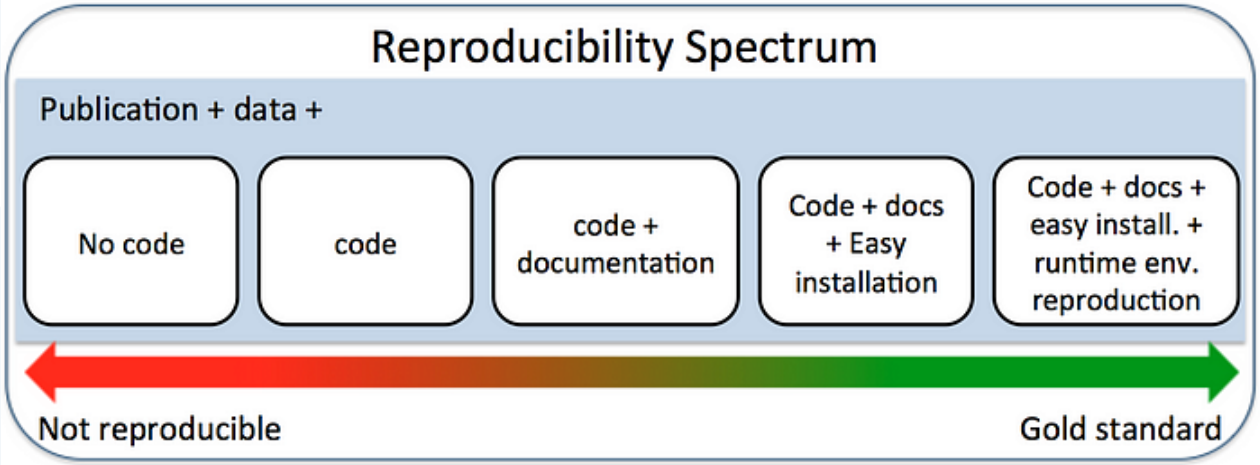
1. Validity (Reproducibility) of scientific findings



American Society of Plant Biologists, Experimental Reproducibility 101: https://plantae.org/experimental-reproducibility-101-part-1/

WHY DO WE NEED OPENNESS IN SCIENCE

1. Validity (Reproducibility) of scientific findings

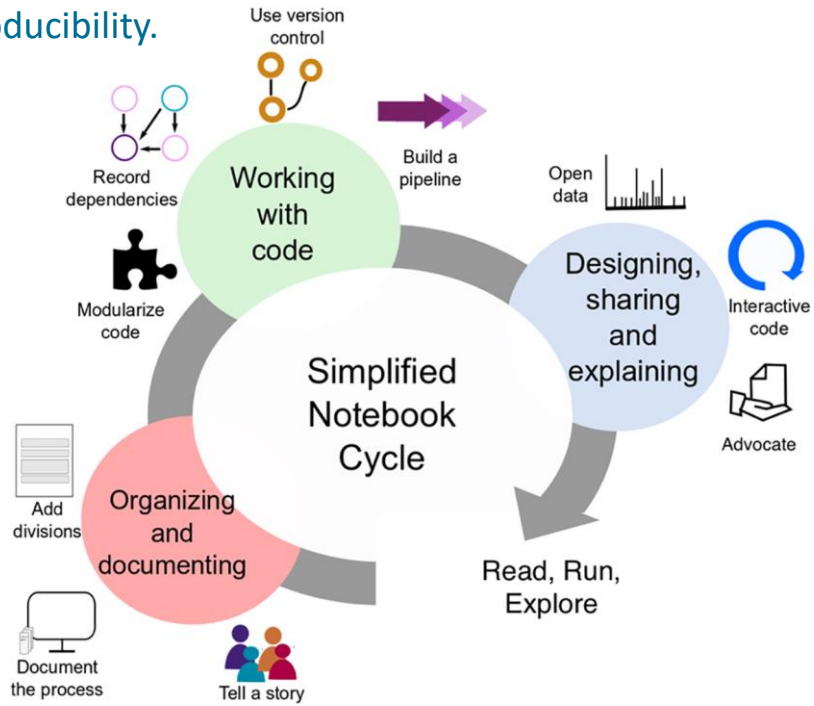


Scientific Data Analysis Pipelines and Reproducibility: https://towardsdatascience.com/scientific-data-analysis-pipelines-and-reproducibility-75ff9df5b4c5

WHY DO WE NEED OPENNESS IN SCIENCE

1. Validity (Reproducibility) of scientific findings

Relevance with Free Software: Free software promotes reproducibility.

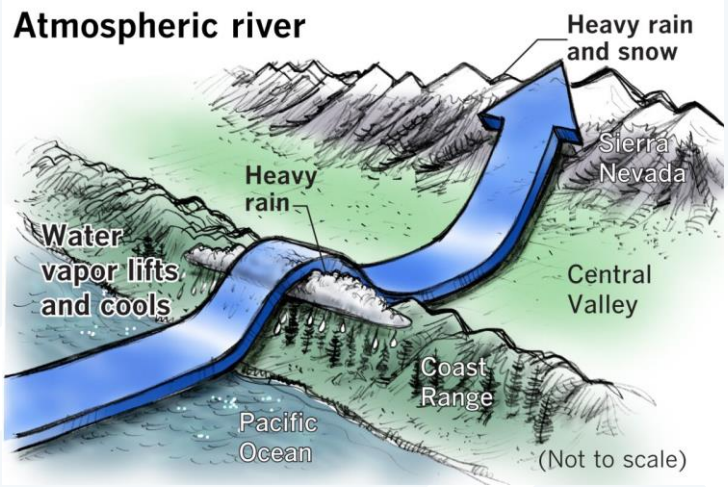


Rule, A., Birmingham, A., Zuniga, C., Altintas, I., Huang, S.C., Knight, R., Moshiri, N., Nguyen, M.H., Rosenthal, S.B., Pérez, F. and Rose, P.W., 2019. Ten simple rules for writing and sharing computational analyses in Jupyter Notebooks. *PLoS computational biology*, *15*(7), p.e1007007.

WHY DO WE NEED OPENNESS IN SCIENCE

2. Increasing complexity and specialization of scientific problems

Example: How can we better understand atmospheric river?

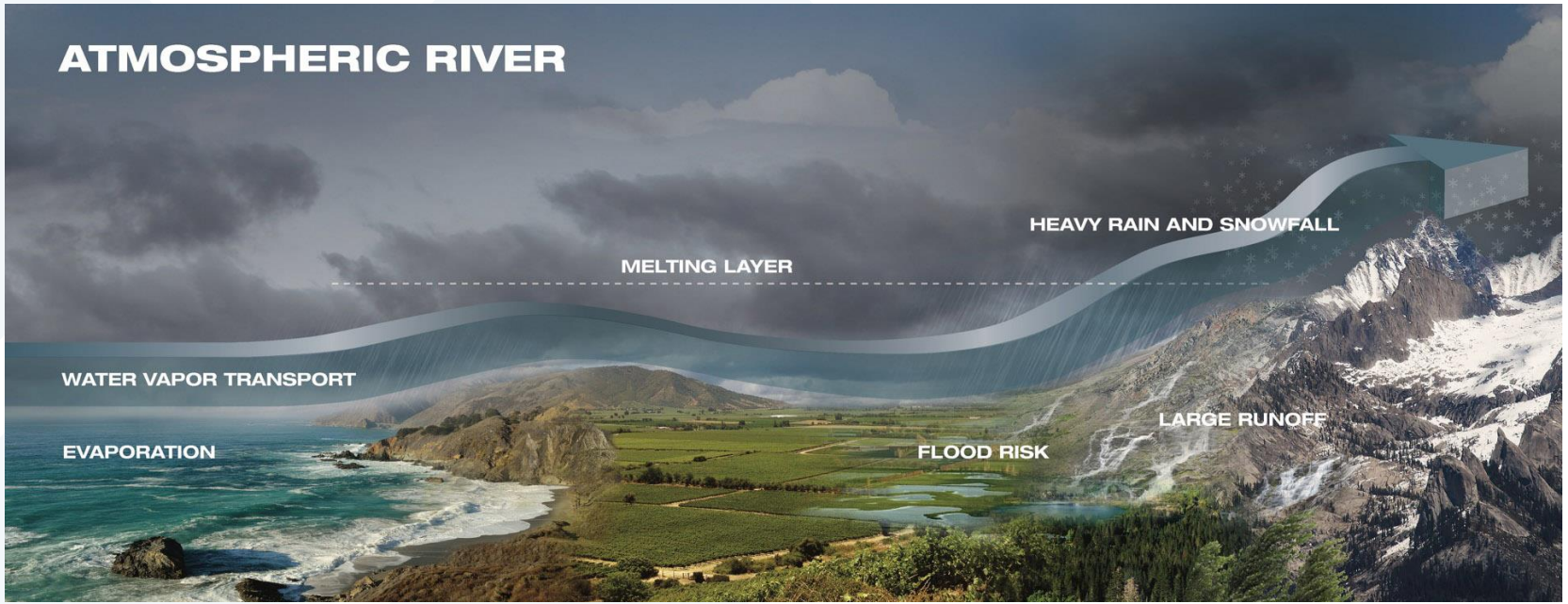


How atmospheric rivers and a bomb cyclone add up to mayhem for California: https://www.latimes.com/california/story/2023-01-03/california-atmospheric-river-bomb-cyclone

WHY DO WE NEED OPENNESS IN SCIENCE

2. Increasing complexity and specialization of scientific problems

Example: How can we better understand atmospheric river?



Rolling on the Science of an Atmospheric River: https://www.jpl.nasa.gov/edu/news/2018/10/4/rolling-on-the-science-of-an-atmospheric-river/

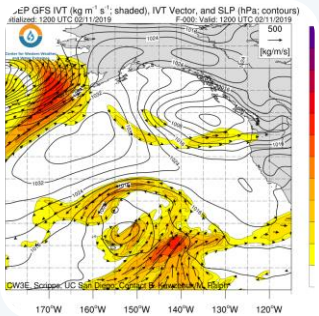
WHY DO WE NEED OPENNESS IN SCIENCE

2. Increasing complexity and specialization of scientific problems

Example: How can we better understand atmospheric river?

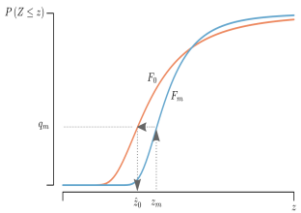
Observation

- Field researcher
- Engineer



Modelling

- Meteorologist
- Computer Scientist



Post-Processing

- Statistician
- Data scientist

Downstream Applications

- Policy maker
- Hydrologist
- Operator
- …

What is a Dropsonde: https://www.eol.ucar.edu/content/what-dropsonde; CW3E: https://cw3e.ucsd.edu

WHY DO WE NEED OPENNESS IN SCIENCE

2. Increasing complexity and specialization of scientific problems

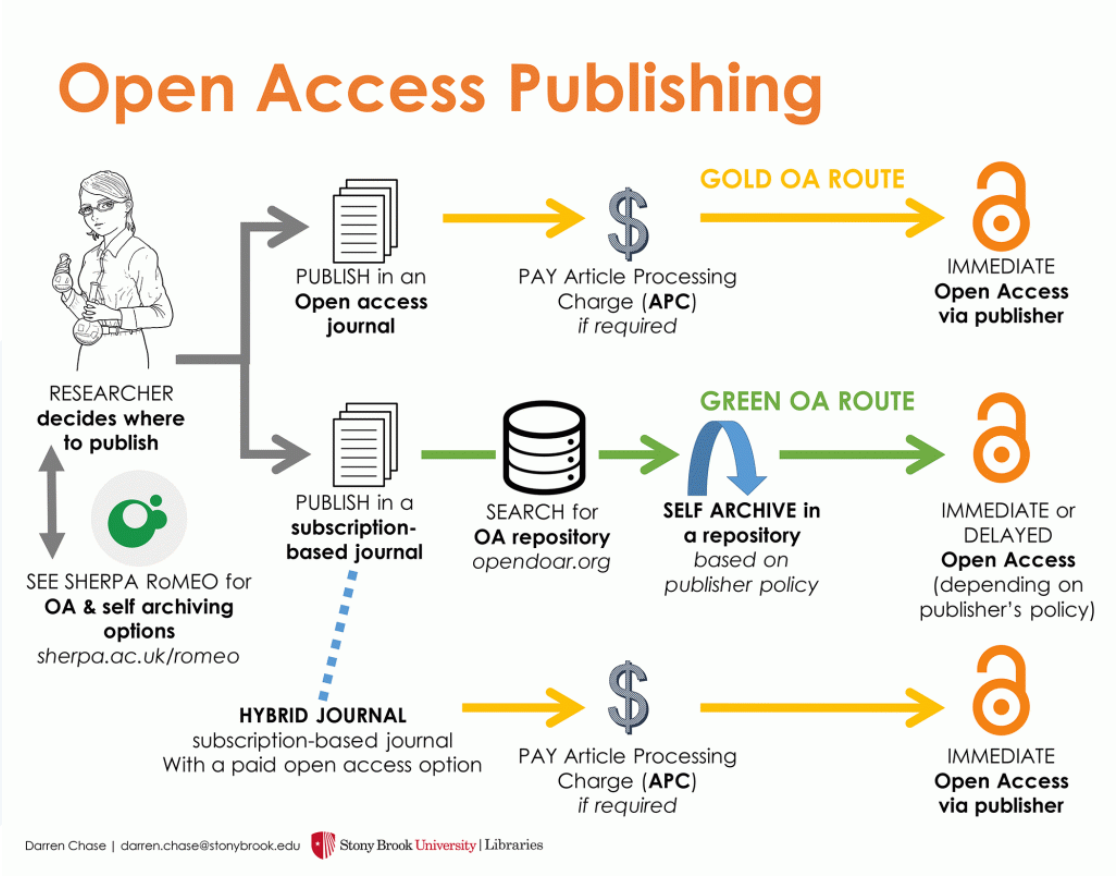
   Example: How can we better understand atmospheric river?

   Relevance with Free Software: Free software facilitate communication and sharing.

OPEN SCIENCE IN A COMPETITIVE ENVIRONMENT

1. Sharing is risky, but the many benefits ultimately outweigh these risks.
    - Visibility
    - Collaboration
2. It drives innovation and prevents over-complexity.
3. Knowledge sharing doesn't necessarily mean giving away all your trade secrets.

OPEN ACCESS IS NOT FREE



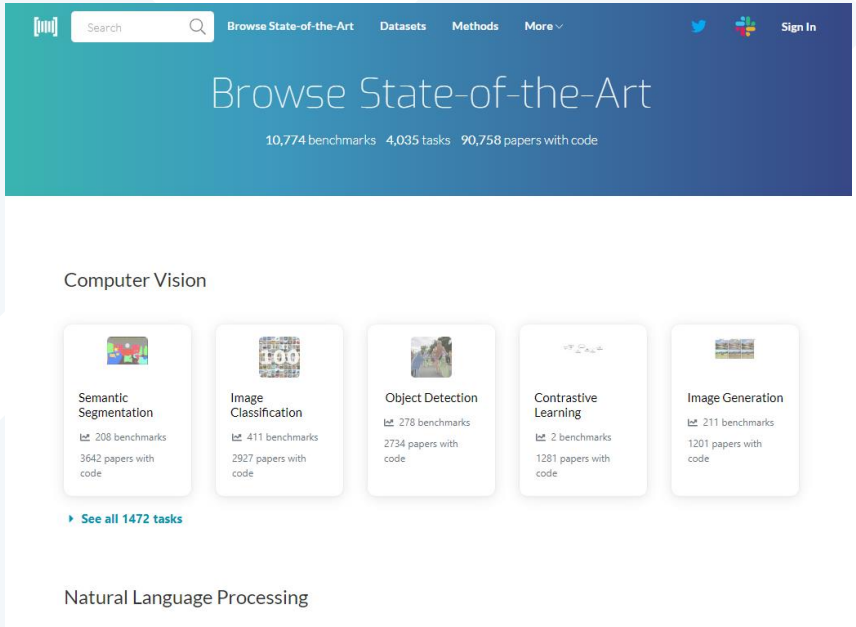
OPEN ACCESS IS NOT FREE

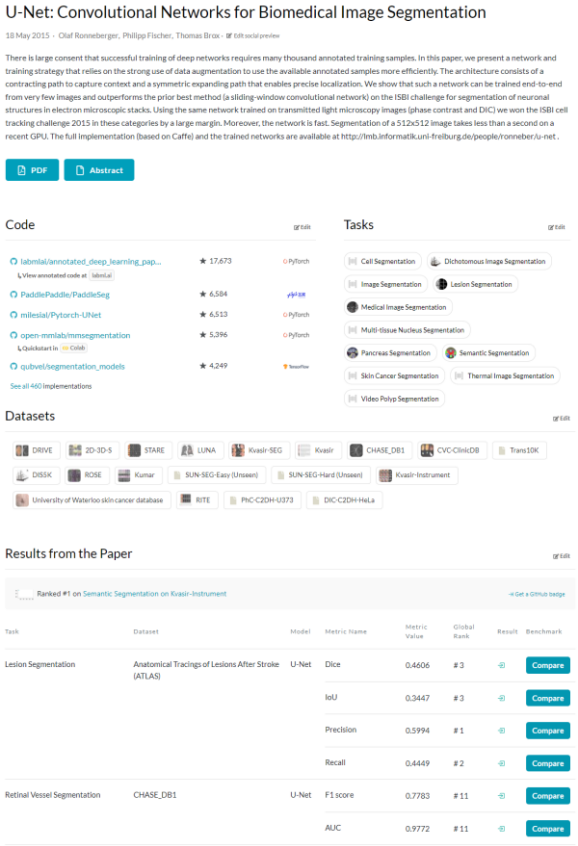
- Predatory journals
- On-going battles with peer review
  - Pre-publication and post-publication review
  - Do you know these phrases? “counterfeit consciousness”’, “profound neural organization” and “colossal information”
- …

FREE SOFTWARE FOR OPEN SCIENCE

1. Reproducibility

- Example: Paper with Code





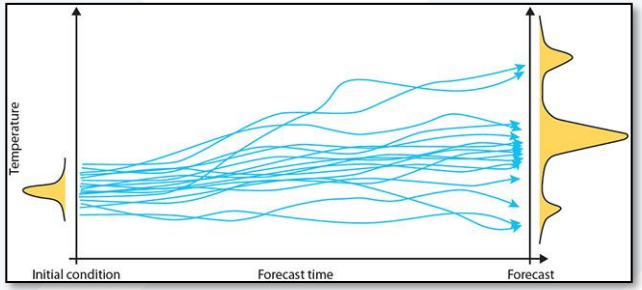
FREE SOFTWARE FOR OPEN SCIENCE

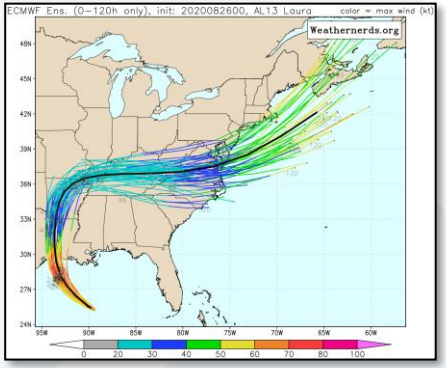
1. Reproducibility
   - Hosting platforms?

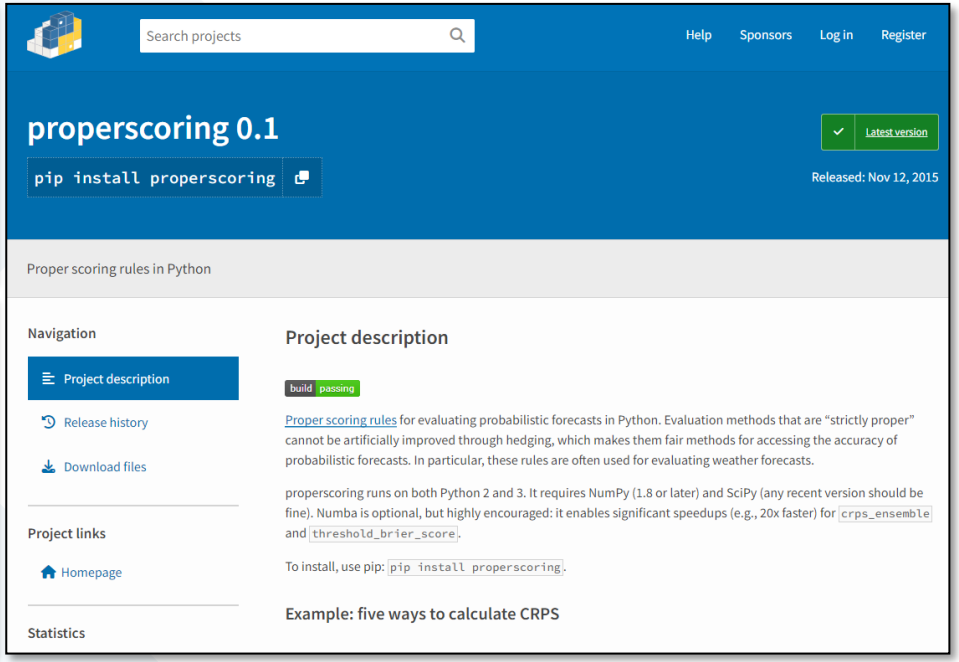
FREE SOFTWARE FOR OPEN SCIENCE

2. Collaboration

- Example: Weather verification







The rapid intensification of hurricane Laura is underway, devastating landfall expected: https://www.severe-weather.eu/tropical-weather/hurricane-laura-landfall-mk/
Twenty-five years of ensemble forecasting: https://www.ecmwf.int/en/about/media-centre/news/2017/twenty-five-years-ensemble-forecasting

FREE SOFTWARE FOR OPEN SCIENCE

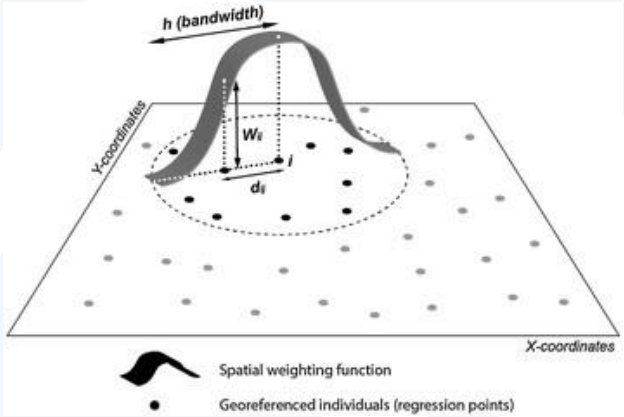
3. Accessibility and Pricing
    - Example: Licensing

TEACHING FREE SOFTWARE FOR ENVIRONMENTAL SCIENCE

Is it possible to only use free / open-source software for environmental science?

TEACHING FREE SOFTWARE FOR ENVIRONMENTAL SCIENCE

An example: Geographically Weighted Regression (GWR)



Add temporal dimension?

Process big data?

Source code?

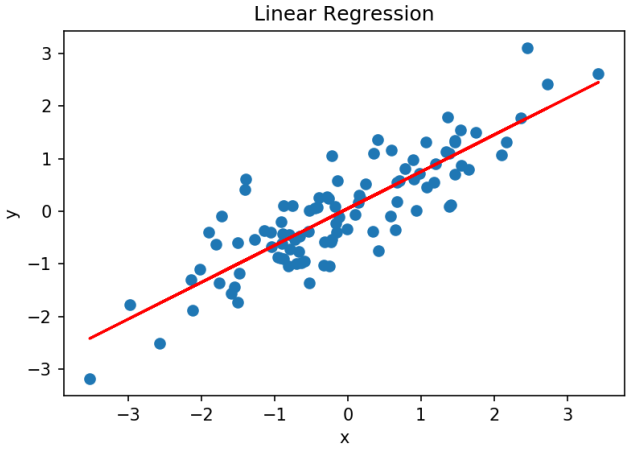
Feuillet, T., Charreire, H., Menai, M., Salze, P., Simon, C., Dugas, J., Hercberg, S., Andreeva, V.A., Enaux, C., Weber, C. and Oppert, J.M., 2015. Spatial heterogeneity of the relationships between environmental characteristics and active commuting: towards a locally varying social ecological model. *International Journal of Health Geographics*, *14*(1), pp.1-14.

TEACHING FREE SOFTWARE FOR ENVIRONMENTAL SCIENCE

- We teach the math first and then teach then how to use the toolbox.

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + ... + b_t X_t + u$$

?



TEACHING FREE SOFTWARE FOR ENVIRONMENTAL SCIENCE

- We don’t need to read the source code for learning purposes.
- Reading and teaching the source code is too time-consuming.
- …

```cpp
// C++ program for implementation
// of Bubble sort
#include <bits/stdc++.h>
using namespace std;

// A function to implement bubble sort
void bubbleSort(int arr[], int n)
{
    int i, j;
    for (i = 0; i < n - 1; i++)

        // Last i elements are already
        // in place
        for (j = 0; j < n - i - 1; j++)
            if (arr[j] > arr[j + 1])
                swap(arr[j], arr[j + 1]);
}

// Function to print an array
void printArray(int arr[], int size)
{
    int i;
    for (i = 0; i < size; i++)
        cout << arr[i] << " ";
    cout << endl;
}

// Driver code
int main()
{
    int arr[] = { 5, 1, 4, 2, 8};
    int N = sizeof(arr) / sizeof(arr[0]);
    bubbleSort(arr, N);
    cout << "Sorted array: \n";
    printArray(arr, N);
    return 0;
}
// This code is contributed by rathbhupendra
```

Bubble Sort Algorithm: https://www.geeksforgeeks.org/bubble-sort/

TEACHING FREE SOFTWARE FOR ENVIRONMENTAL SCIENCE

1. Choose the right software
2. Provide context
3. Start with examples and hands-on tutorial
4. Guide them to investigate the implementation
5. Collaborative learning
6. Make it fun

WHAT CAN THE DEV COMMUNITY DO

1. Be patient
2. Create educational resources
3. Collaborating with educators


Free software is not just about technology;
it’s about values.
Education is also not just about occupation;
its about critical thinking and being creative.