Project Pythia

An Open Source Educational Resource for Geoscience Data Analysis

John Clyne[1], Matt Long[1], Ryan May[2], Kevin Paul[1], Brian Rose[3] and Kevin Tyle[3]

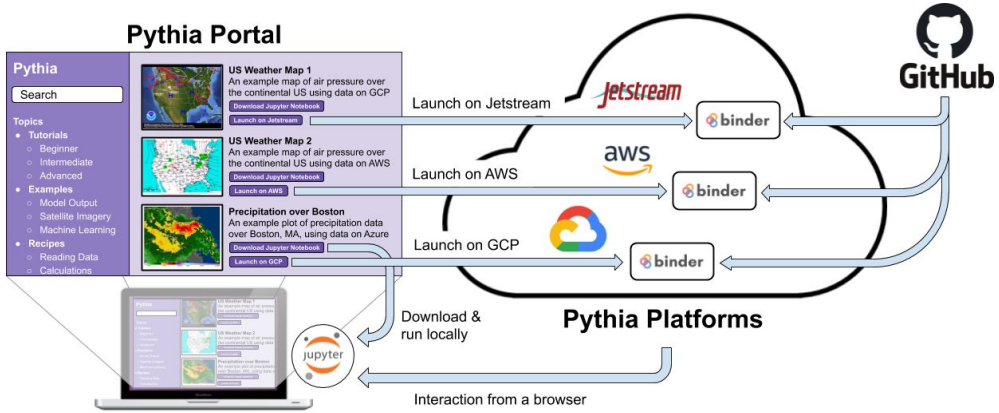
[1] National Center for Atmospheric Research
[2] Unidata
[3] University at Albany

ARM/ASR Open Science Virtual Workshop
May 10, 2022

This material is based upon work supported by the National Science Foundation under Grant Nos. 2026863 and 2026899. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Project Pythia

A Community Learning Resource for Geoscientists



Aspiration goal: Be the goto resource for learning the *Scientific Python Ecosystem*

- ★ Geoscience focused
- ★ From beginner to the power user
- ★ Tutorials, videos, examples, on-line courses, and sample data
- ★ Community owned!

This work supported through the National Science Foundation Award #2026899.

The *Pythia Portal*

PROJECT PYTHIA

Training resources

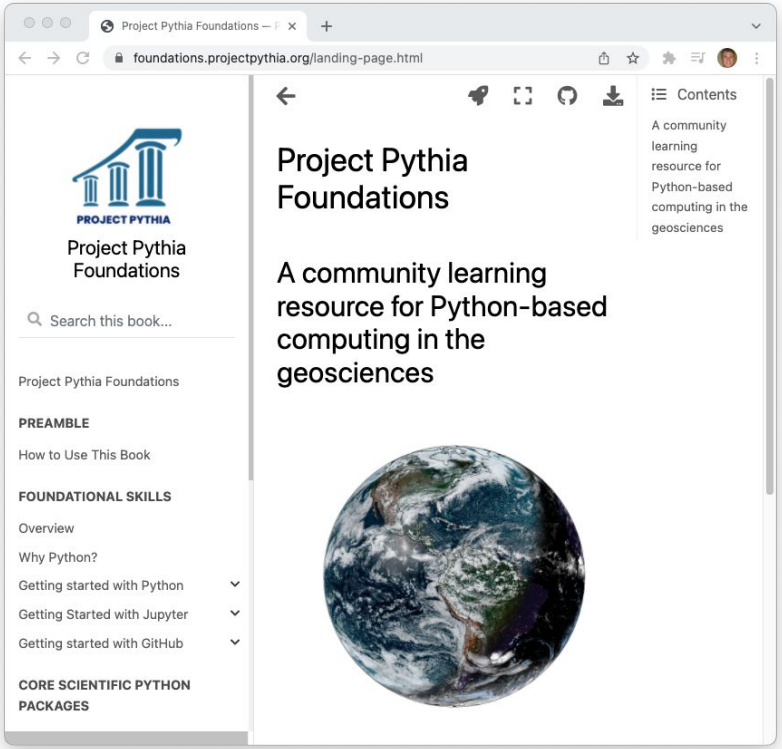
- *Binderized* Jupyter Notebooks
  - Download, share, run
- Sample geoscience data
- Tutorials (on-line and webinars)
- Communication forums
- Many, many links to external content

Sample content

- Introductory - ***Pythia Foundations***
  - NumPy & Xarray
  - Conda
  - Git & GitHub
  - Jupyter Notebooks
- Geoscience-focused packages, e.g.
  - MetPy
  - GeoCAT
- Scalable workflows
  - Dask
  - Using Cloud resources

Pythia Foundations

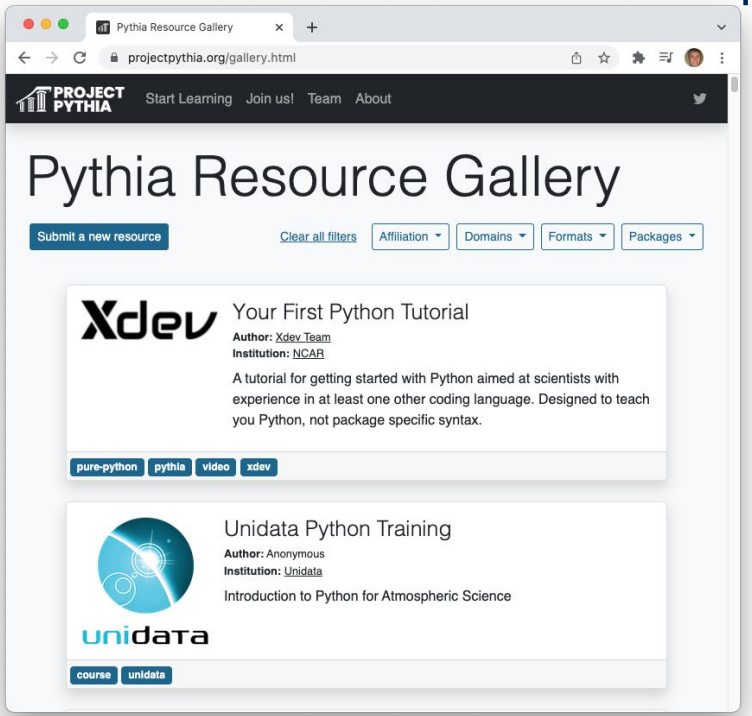
What every geoscientist should know



Pythia Resource Gallery

Links to internal *and* external resources

- Search
- Filter
- Contribute!

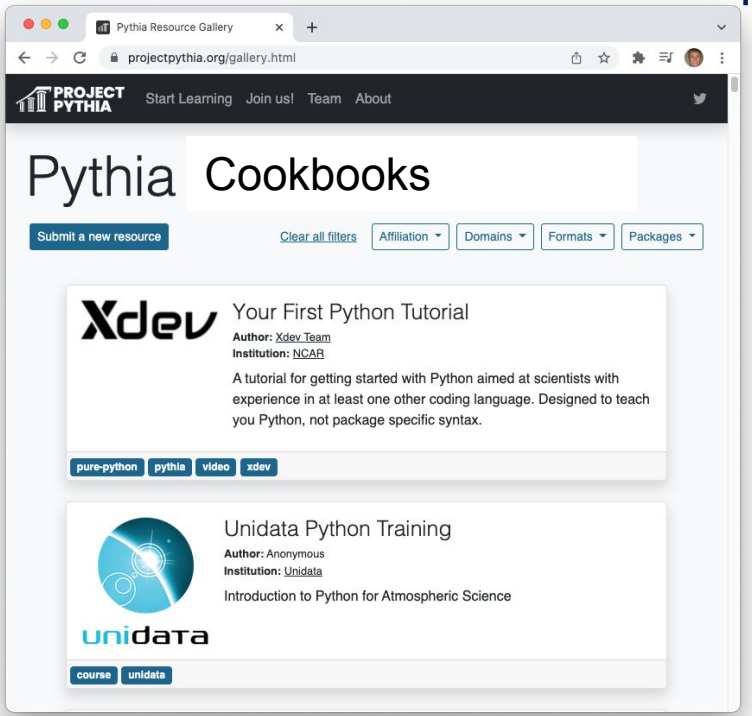


Coming soon: Pythia Cookbooks

A collection of short recipes that demonstrate how to perform a particular analysis task.

- E.g. Using the Python ARM Radar Toolkit (Py-ART) to plot weather radar data

- Search
- Filter
- Contribute!



Next up for Pythia: The *Pythia Platform*

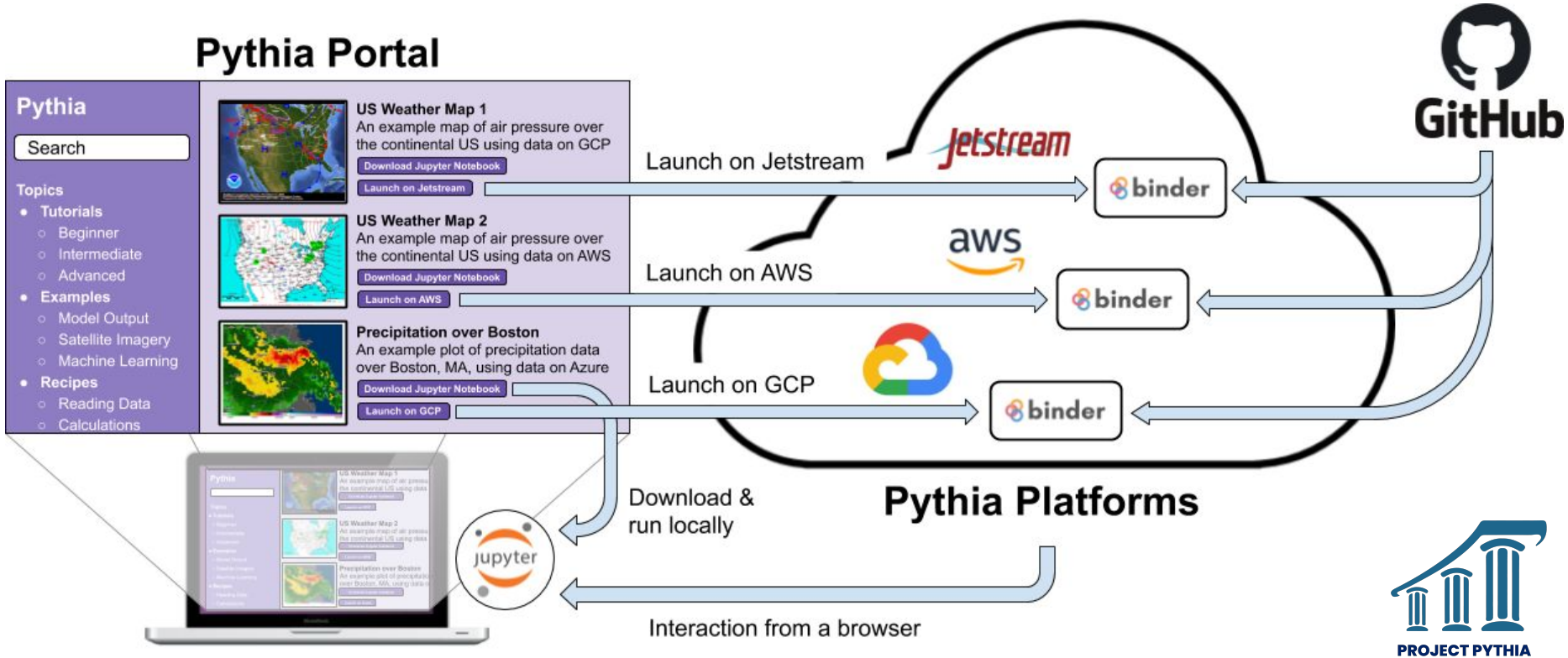
Binder-like utilities that support interactive cloud execution environments for each Jupyter Notebook

- Example notebooks on Pythia Portal will be executable on cloud resources
- Scalability beyond *MyBinder*
- Authenticated access (when needed)
- Redeployable outside of Project Pythia

Targeted Cloud resource providers:

- AWS, GCP, Azure, XSEDE Jetstream2, and HPC (e.g. Casper/Cheyenne)

Pythia Portal + Pythia Platforms



Open Development

Project Pythia is a community-owned resource and follows an *Open Development* model. The user community is expected to contribute by:

1. Providing feedback on Project Pythia resources
2. Helping identify and prioritize content needs
3. Helping develop new content or identify existing content for inclusion
4. Responding to questions from other users
5. Reporting or correcting problems
6. … and more

All Project Pythia-developed content is hosted on GitHub

Summary

Project Pythia is a community-owned, web-accessible educational resource for helping geoscientists at all levels of their career become proficient with the Scientific Python Ecosystem (SPE)

Three pillars of the SPE: The Python language, GitHub, Jupyter

Community engagement is essential for long term sustainability of Pythia

Please get involved!!!

https://projectpythia.org

Acknowledgements

Pythia community of contributors, especially numerous technical staff doing the heavy lifting at NCAR, Unidata, University at Albany

Pangeo community

NSF Earth Cube program (award #2026899, 2026863)