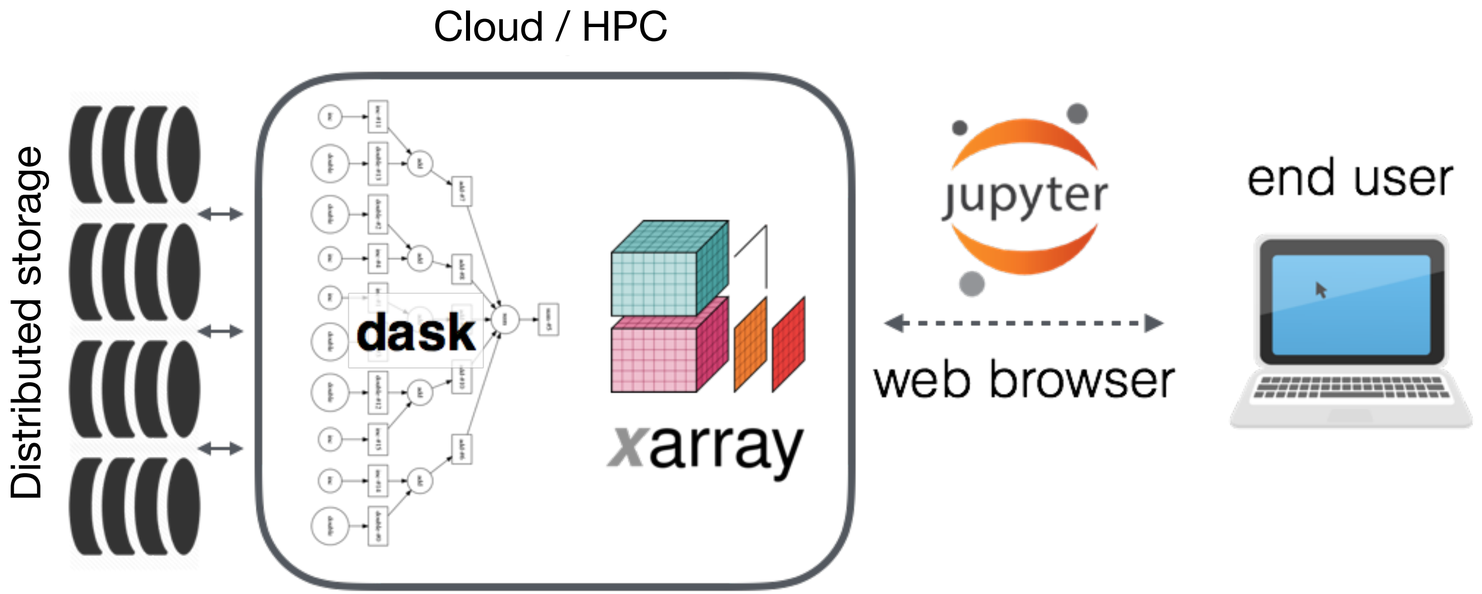


intake-esm Tutorial

Aaron Spring (MPI-M)

PANGEO

What will happen today?

- 10’ intro: me 🗣️ you 🍕
- 10’ tutorial: `tutorial.ipynb` 💫
- 20-100’ hands-on: `tasks.ipynb` 👩‍💻

What is Pangeo?

"A community platform for Big Data geoscience"

- Open Community
- Open Source Software
- Open Source Infrastructure

Pangeo Community

Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE

NSF

EARTHCUBE

NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Met Office

NASA

USGS

ECMWF

GEOMAR

UNIVERSITY of WASHINGTON
eScience Institute

BIDS
BERKELEY INSTITUTE
FOR DATA SCIENCE

jupyter

cnes
CENTRE NATIONAL D'ÉTUDES SPATIALES

CSIRO

ANACONDA
Powered by Continuum Analytics

JUPITER

element84

developmentSEED

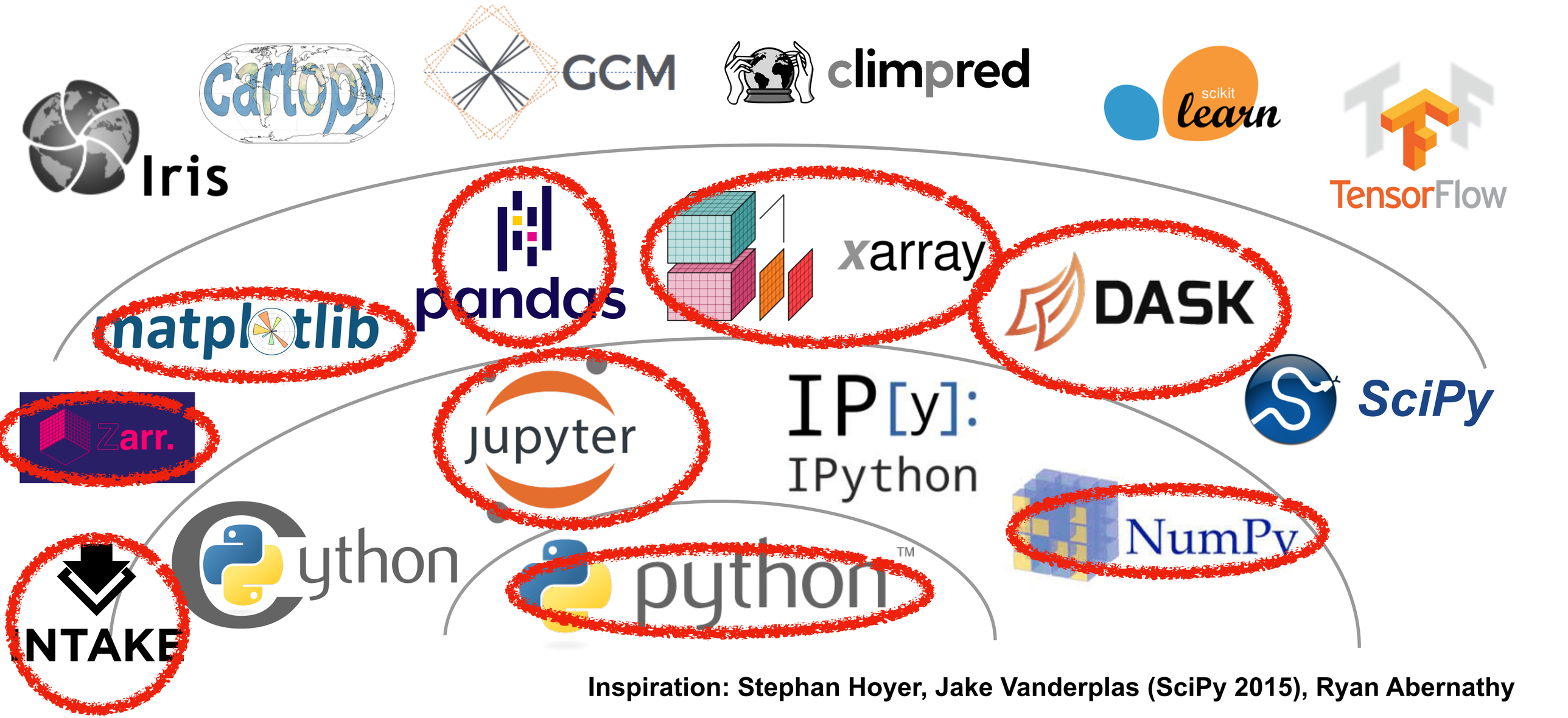
RHODIUM GROUP

CLIMACELL
Weather Revealed

http://pangeo.io

PANGEO

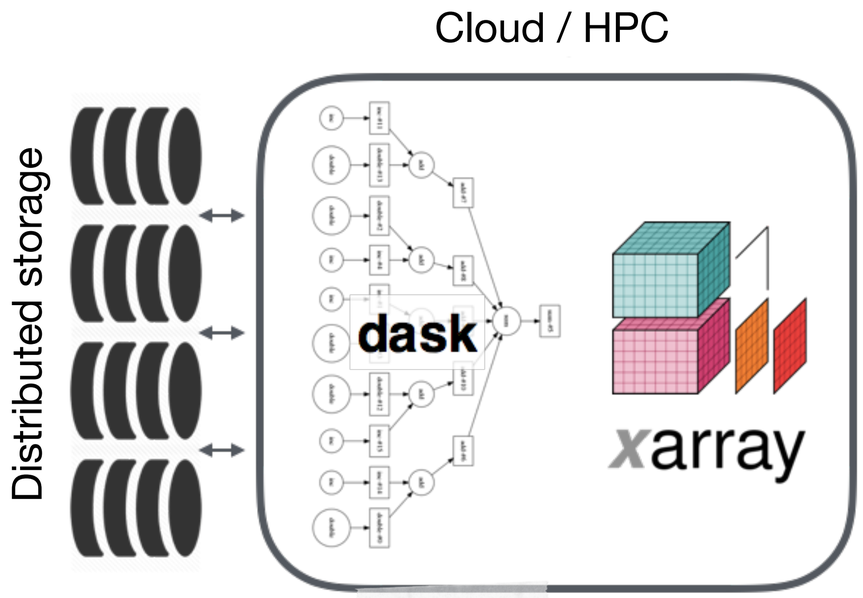
Pangeo Software Ecosystem



Inspiration: Stephan Hoyer, Jake Vanderplas (SciPy 2015), Ryan Abernathy

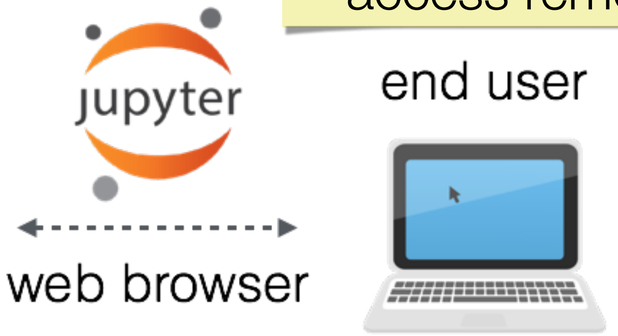
HPC Architecture

"Analysis Ready Data" stored on distributed storage.



Parallel computing system allows users deploy clusters of compute nodes for data processing.

Dask tells the nodes what to do.

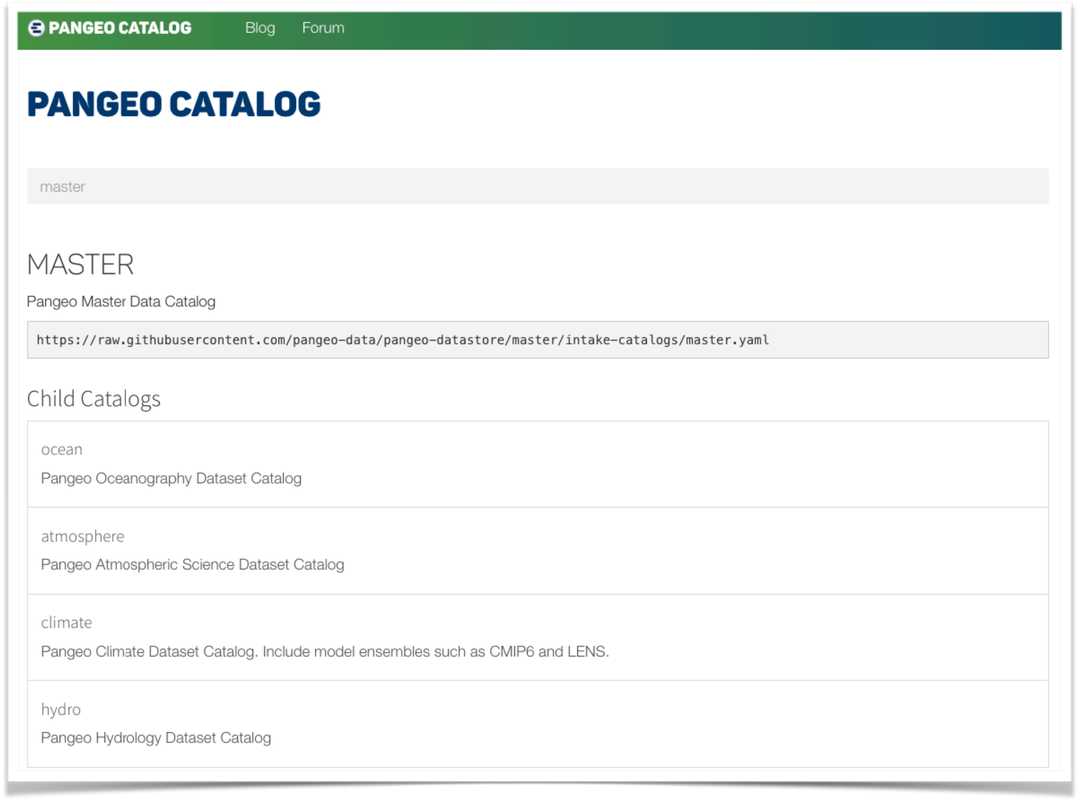


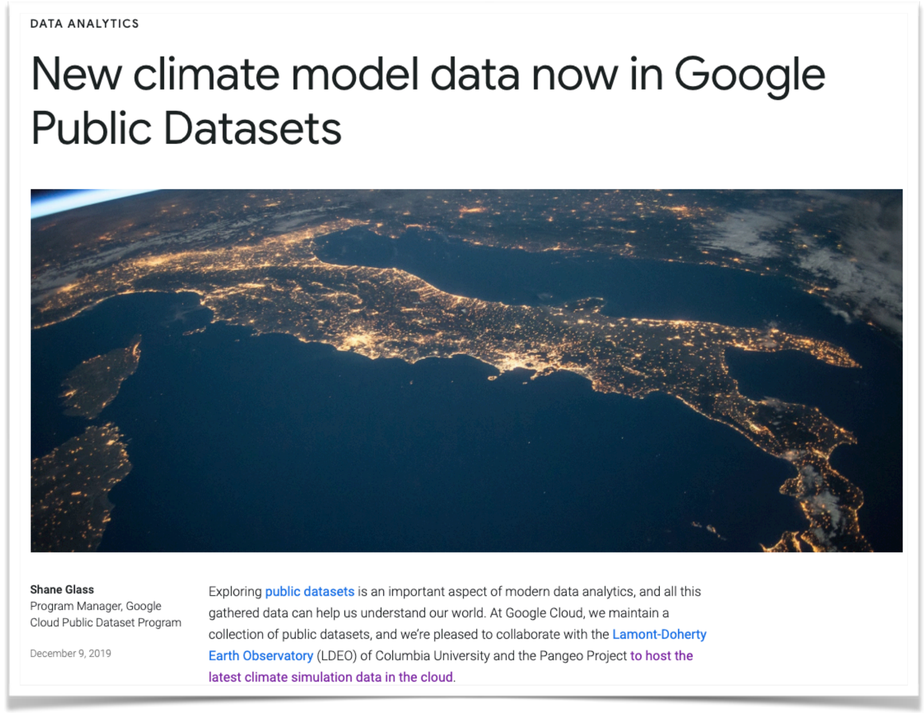
Jupyter for interactive access remote systems

Xarray provides data structures and intuitive interface for interacting with datasets

Pangeo Cloud Data Catalog

catalog.pangeo.io





launch binder

Culture of sharing

- Together we achieve more.
- Currently we make reproducibility unnecessary hard.
  - I try my best: https://github.com/aaronspring/Spring_etal_2020_Code
- Let's put code examples out there:
  - `gitlab.dkrz.de`
  - `github.com`
  - `gist.github.com`

intake-esm

- Developed at NCAR, mostly by Anderson Banihirwe (**andersy005**)
- https://intake-esm.readthedocs.io/en/latest/
- Taking the pain out of data access: link to wiki.mpimet
- Search and load ESM output:
  - Query in `pandas.DataFrames`
  - Load data with `dask` into `xarray`

Live demo

tutorial.ipynb

launch binder

Try intake-esm yourself!

```
git clone https://gitlab.dkrz.de/
  m300524/lunchbytes_intake-esm
```

OR

launch binder