XARRAY IN CLIMATE SCIENCE

Part 1: Introduction to Xarray

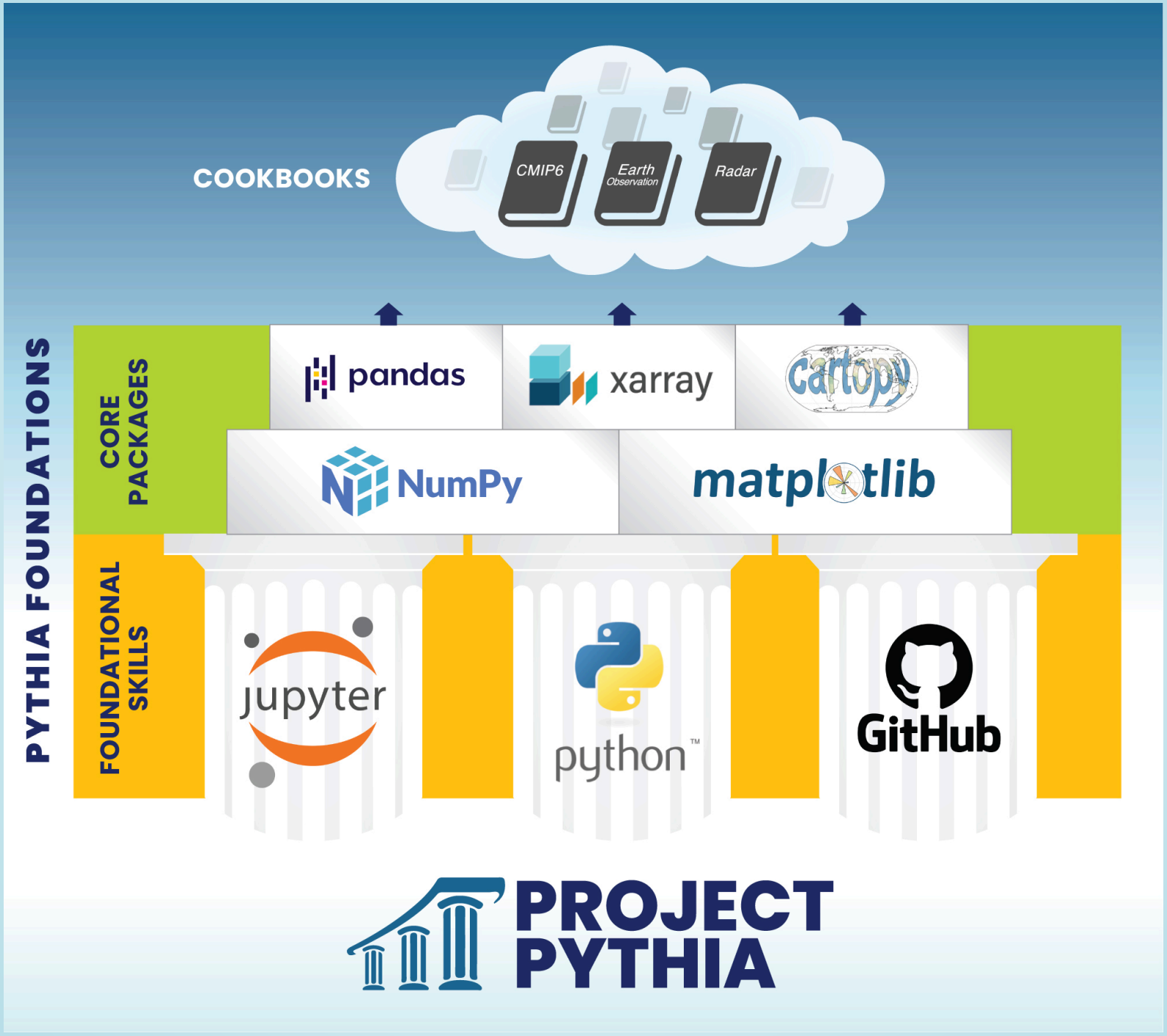
(based on the Project Pythia Foundations book)

STRUCTURE OF THIS MINI-SERIES

- Part 1: Introduction to Xarray
- Part 2: Computations and Masks with Xarray
- Part 3: ???

THE BIGGER PICTURE: PYTHON-BASED COMPUTING IN THE GEOSCIENCES

Pythia Foundations: A community learning resource for Python-based computing in the geosciences


COOKBOOKS
CMIP6
Earth Observation
Radar
PYTHIA FOUNDATIONS
CORE PACKAGES
pandas
xarray
Cartopy
NumPy
matplotlib
FOUNDATIONAL SKILLS
jupyter
python™
GitHub
PROJECT PYTHIA

STRUCTURE OF THIS TALK

- `[2min]` The problem of multidimensional array data in climate science
- `[5min]` Why Xarray?
- Following the Pythia Foundations "Introduction to Xarray" notebook

MULTIDIMENSIONAL ARRAY DATA IN CLIMATE SCIENCE

Copernicus Marine Ocean Viewer

TECHNICAL CHALLENGES

TECHNICAL CHALLENGES

- Lots of metadata
  - -> Needs to be handled carefully for correct observations

TECHNICAL CHALLENGES

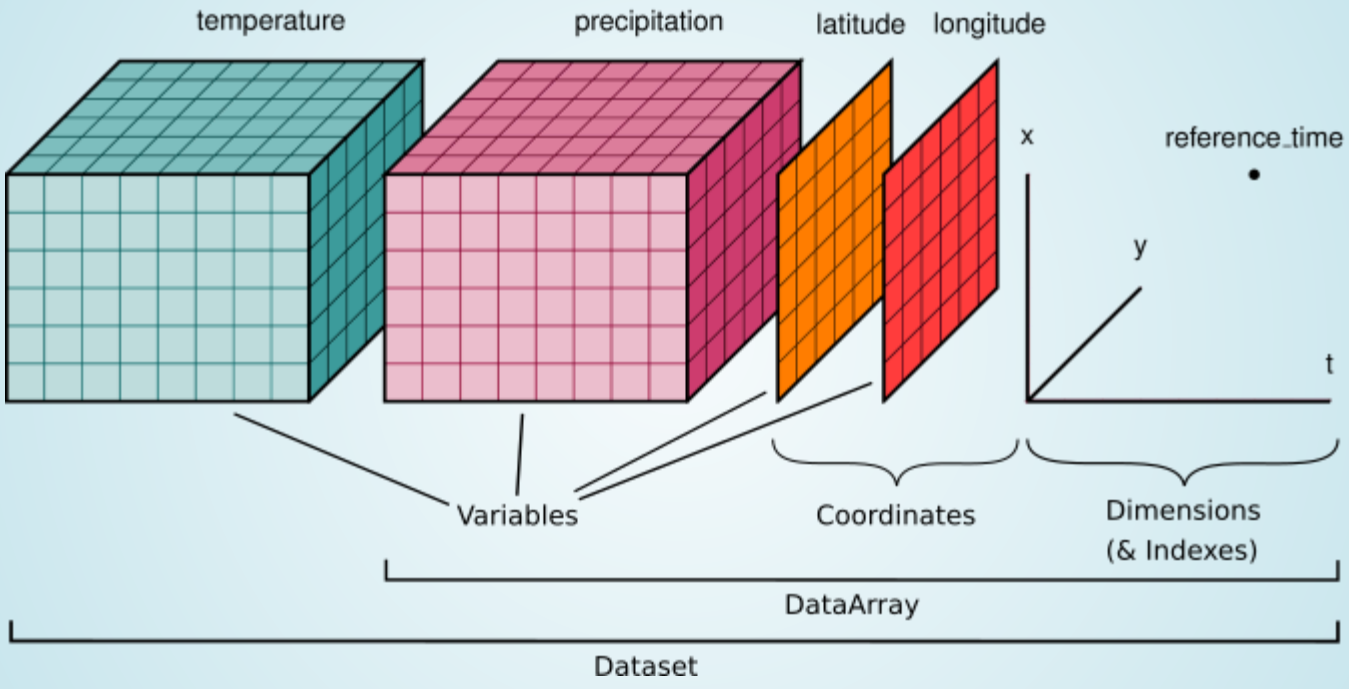
- Lots of metadata
  - -> Needs to be handled carefully for correct observations
  - -> Working between datasets needs alignment of the raw data

TECHNICAL CHALLENGES

- Lots of metadata
  - -> Needs to be handled carefully for correct observations
  - -> Working between datasets needs alignment of the raw data
- Big data
  - -> How can we work with data that can't fit in memory?

WHY XARRAY?

Xarray introduces labels in the form of dimensions, coordinates and attributes on top of raw NumPy-like multidimensional arrays, which allows for a more intuitive, more concise, and less error-prone developer experience.


temperature
precipitation
latitude
longitude
x
reference_time
y
t
Variables
Coordinates
Dimensions
(& Indexes)
DataArray
Dataset

LET'S LOOK AT THE PROJECT PYTHIA XARRAY INTRO!

Project Pythia: Introduction to Xarray

ADDITIONAL RESOURCES

- Xarray Documentation
- Xarray Tutorial