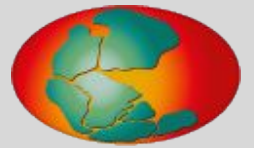# PANGAEA
# Community Workshop 2022

## Introduction to PangaeaPy

*Kathrin Riemann-Campe and Michael Oellermann*

# Before we start…
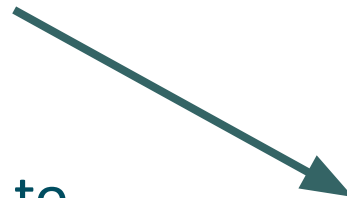
Launch the Python cloud environment!

*GitHub:*

https://github.com/pangaea-data-publisher/community-workshop-material

→ click on 2nd link "LaunchBinder"

…it may take a short while to load…

# Before we start…
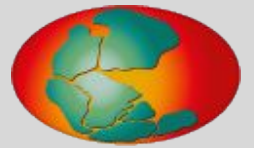
## Run KeepBinderActive.ipynb script to prevent timeout

# Why PangaeaPy?

Direct data queries and retrieval via Python

Easy bulk data download

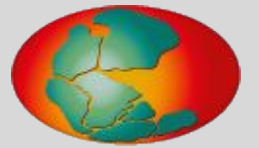Metadata search and analysis

Enables automated workflows and data mining

Clear record of data retrieval workflow and data sources

Credit: Gregor Cesnar

# Getting started

## At work/homeoffice

Install PangaeaPy in your environment

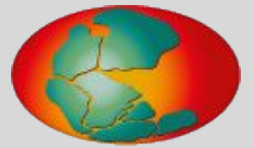"pip install pangaeapy"

Import packages

import pangaeapy as pan

or

from pangaeapy.panquery import PanQuery
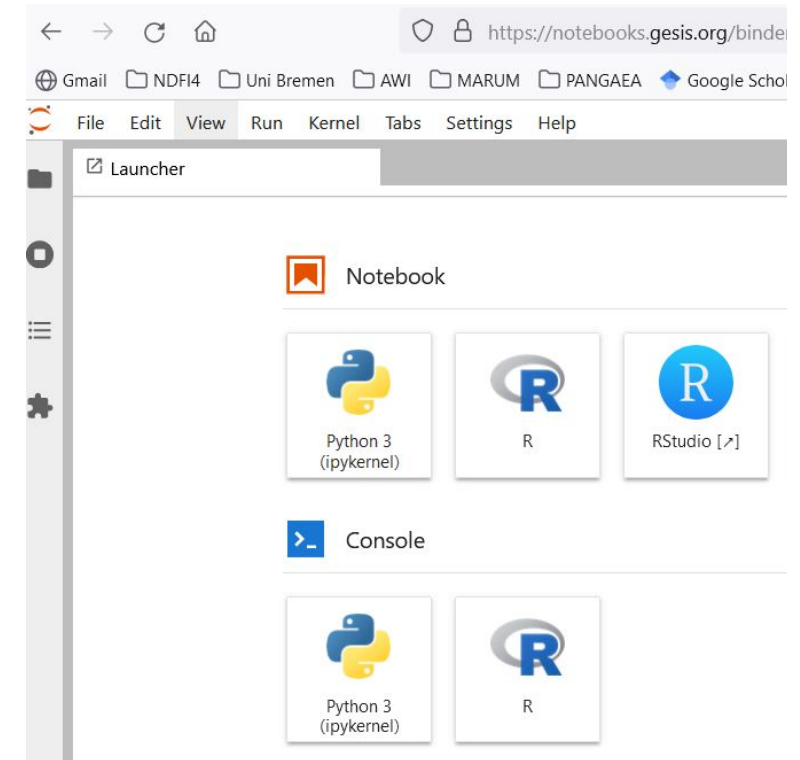
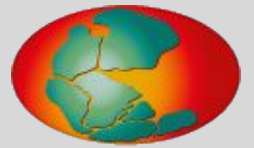from pangaeapy.pandataset import PanDataSet

# Getting started

## For our practice today

- Work in cloud environment (MyBinder)

- Linked to PANGAEA`s GitHub account

- JupyterLab, R or RStudio

  → cannot save files but download

# Working in Jupyter notebooks

Run cells with Shift + Enter or clicking play button



Linked GitHub File repository

Code and Markdown

# PangaeaPy commands
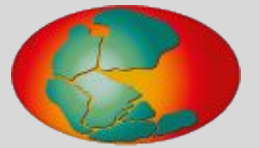
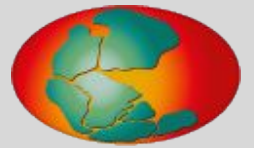| Function | Attributes |
|---|---|
| pan.PanQuery('search term') | totalcount, result |
| pan.PanAuthor('author name') | lastname, firstname, ORCID |
| pan.PanEvent('label') | label, lat, long, elevation, date, device |
| pan.PanParam | id, name, param_type, unit, PI |

# PangaeaPy commands
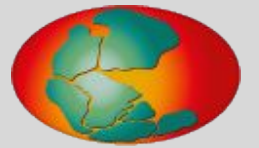
| Function | Attributes |
|---|---|
| pan.PanDataSet('PANGAEA ID') | data, id, uri, doi, title, abstract, year, authors, citation, parameters, events, min-/maxtimeextent, children, etc. |

# Let`s practice

KeepBinderActive running

Navigate to Python/Examples

Start *"getting_started_pangaeapy.ipynb"*

To switch to RStudio later on:

Turn off simple view →

Launcher Tab → RStudio