Climate data processing for climate resilience

Tajikistan and Kyrgyzstan

Data access, processing and methodological concepts

Webinar
17. - 27. 11.2020

DAY 05
The Big Data Problem

Objectives of Day 03

Date: 22.11
Presentation: Day05_BigData

https://github.com/nilshempelmann/climdatatutorial

Recommended: https://github.com/nci-training/notebooks-python

Objectives:

- Wrap up of last week
-

Growing Amount of Data (Example Climate Model data)

CMIP = Coupled Model Inter-comparison Project
IPCC = Intergovernmental Panel of Climate Change
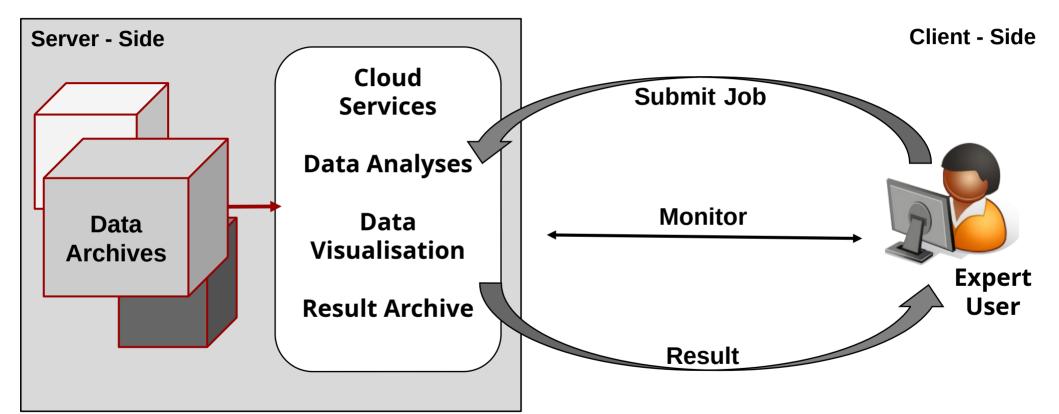**ESGF = Earth System Grid Federation**

Internet Bandwidth
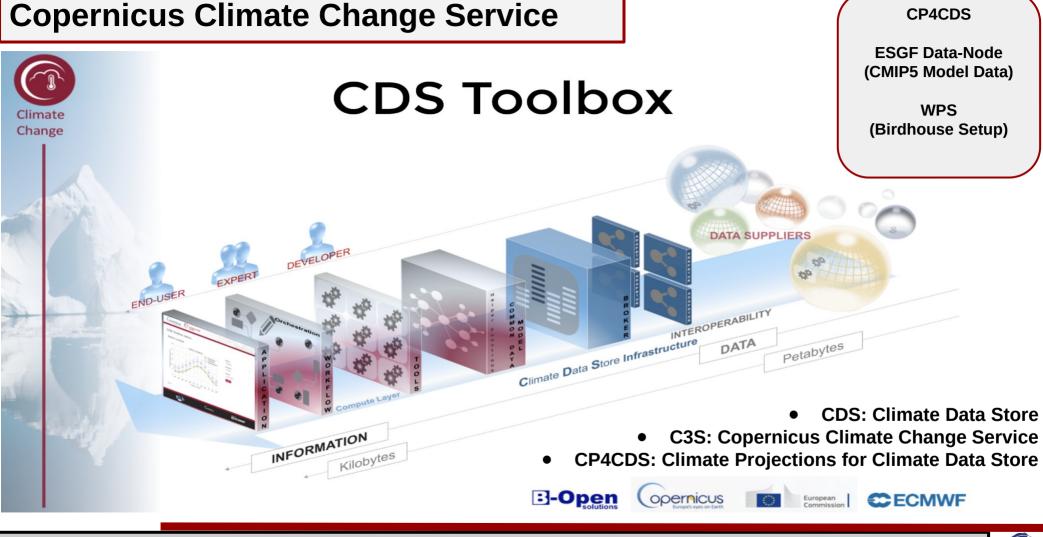
Web Processing Service

high performance environment

(low internet bandwidth)

Server - Side

Client - Side

Data Archives

Cloud Services

Data Analyses

Data Visualisation

Result Archive

Submit Job

Monitor

Result

Expert User

Copernicus Climate Change Service

CDS Toolbox

- **CDS: Climate Data Store**
- **C3S: Copernicus Climate Change Service**
- **CP4CDS: Climate Projections for Climate Data Store**
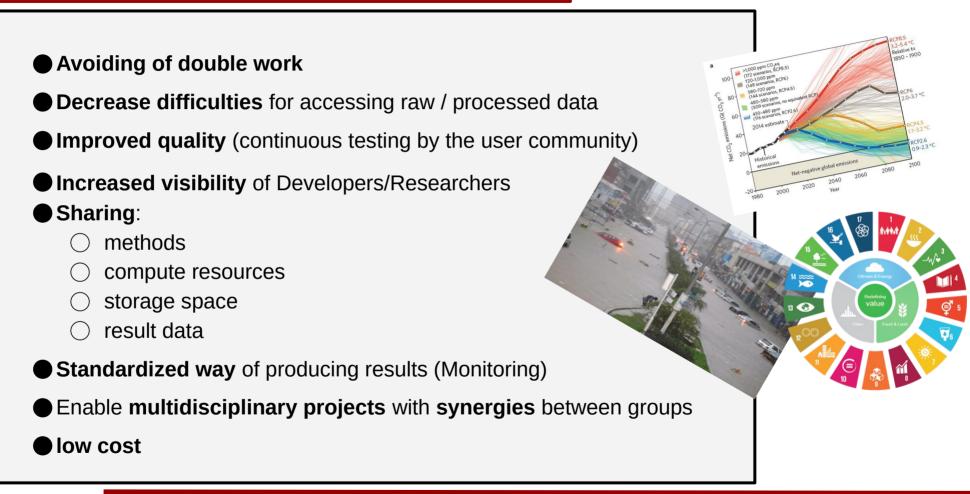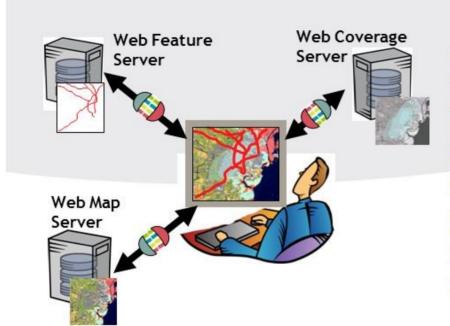
Advantage of Server-Side Services

- **Avoiding of double work**

- **Decrease difficulties** for accessing raw / processed data

- **Improved quality** (continuous testing by the user community)

- **Increased visibility** of Developers/Researchers
- **Sharing**:
  - methods
  - compute resources
  - storage space
  - result data

- **Standardized way** of producing results (Monitoring)

- Enable **multidisciplinary projects** with **synergies** between groups

- **low cost**

OGC Standard

The geospatial web is enabled by OGC standards:

Web Feature Server

Web Coverage Server

Web Map Server

Web Map Service (WMS)
Web Map Tile Service (WMTS)
Web Feature Service (WFS)
Web Coverage Service (WCS)
Catalogue (CSW)
Geography Markup Language (GML)
KML
Others…

Relevant to geospatial applications: Critical Infrastructure, Emergency Management, Weather, Climate, Homeland Security, Defense & Intelligence, Oceans Science, etc

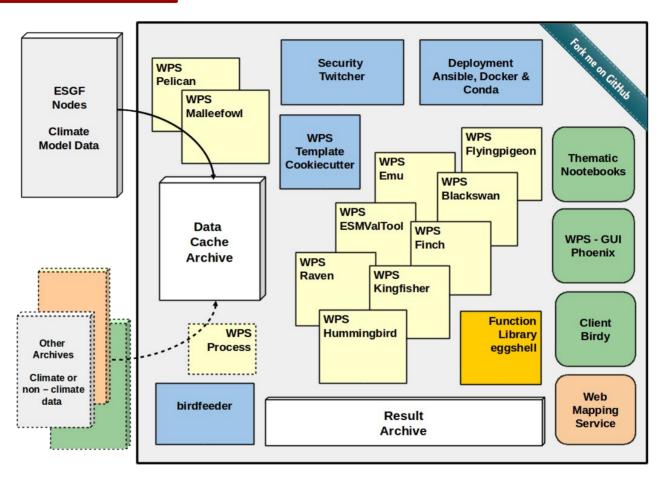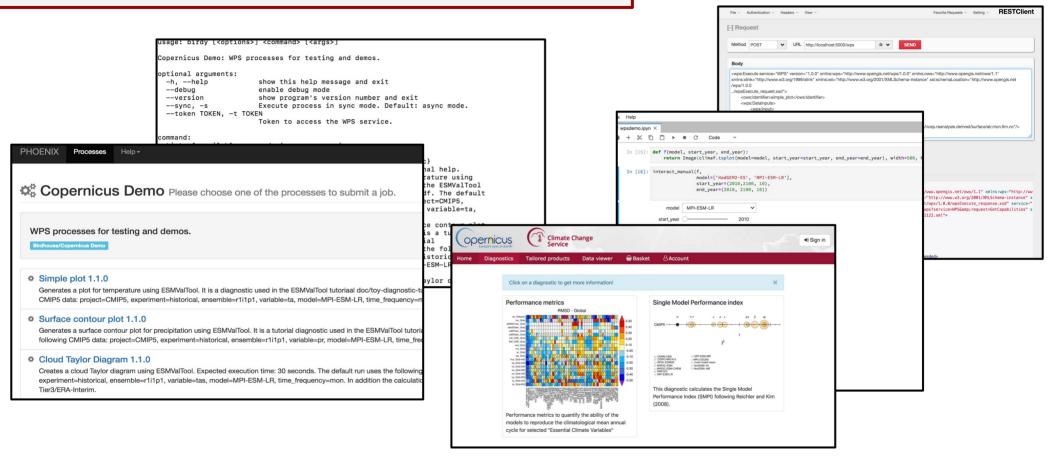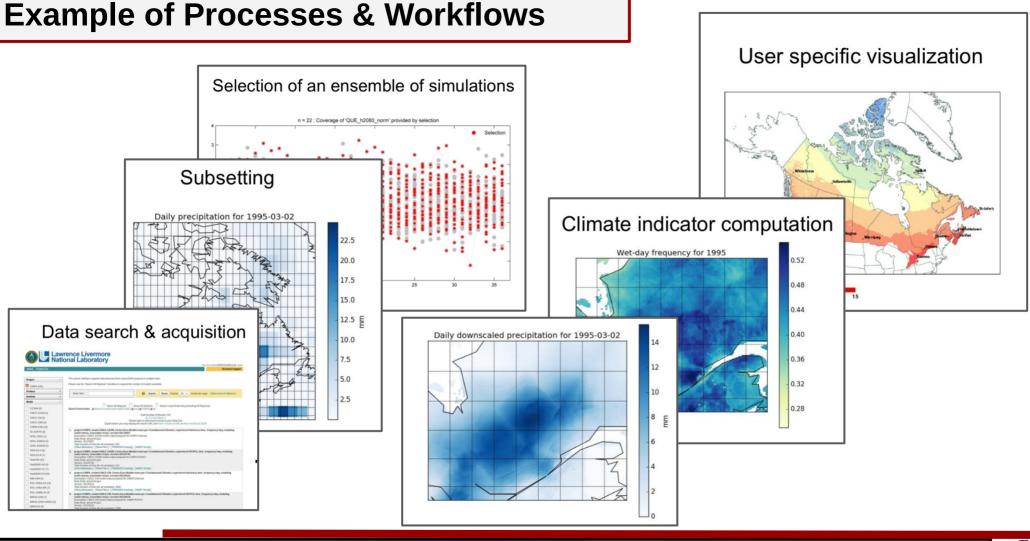*Osservare per prevedere, prevedere per prevenire*

Birdhouse Ecosystem

Open Source Software compartements to set up Open Standared based Web Processing Services (WPS)

ESGF Nodes

Climate Model Data

WPS Pelican

WPS Malleefowl

Security Twitcher

Deployment Ansible, Docker & Conda

Fork me on GitHub

WPS Template Cookiecutter

WPS Emu

WPS Flyingpigeon

WPS Blackswan

WPS ESMValTool

WPS Finch

Thematic Nootebooks

Data Cache Archive

WPS Raven

WPS Kingfisher

WPS - GUI Phoenix

Other Archives

Climate or non – climate data

WPS Process

WPS Hummingbird

Function Library eggshell

Client Birdy

birdfeeder

Result Archive

Web Mapping Service

Client Side: Many Views on WPS

Example of Processes & Workflows

Selection of an ensemble of simulations

Subsetting

User specific visualization

Climate indicator computation

Data search & acquisition
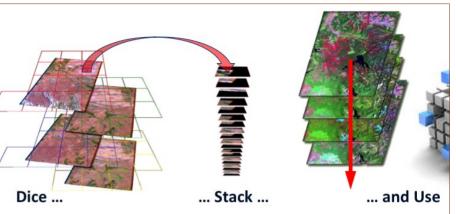
Daily downscaled precipitation for 1995-03-02

Open Data Cube

What are (EO) Data Cubes?

- **Data Cube** = Spatially aligned pixels ready for analysis
- **Proven concept** by Geoscience Australia (GA), the Australian Science Agency (CSIRO) and a super-computer facility (NCI)
- **Shift in Paradigm** ... Pixels vs Files
- **Analysis Ready Data (ARD) ...** Reduce processing burden on users
- Supports **integration** of multiple datasets.
- **Multi-platform** … PC, HPC, Cloud

Open Source Software
https://github.com/opendatacube

TIME

Dice ...   ... Stack ...   ... and Use

https://github.com/opendatacube

Including example notebooks

Open Data Cube Initiative

https://www.opendatacube.org/

3 — Operational
6 — Under Development
31 — Under Review or Expressed Interest

Other processing services

Hands ON:

Climate indice with WPS