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

info@nilshempelmann.de

Objectives of Day 05

Date: 23.11
Presentation: Day05_BigData

https://github.com/nilshempelmann/climdatatutorial

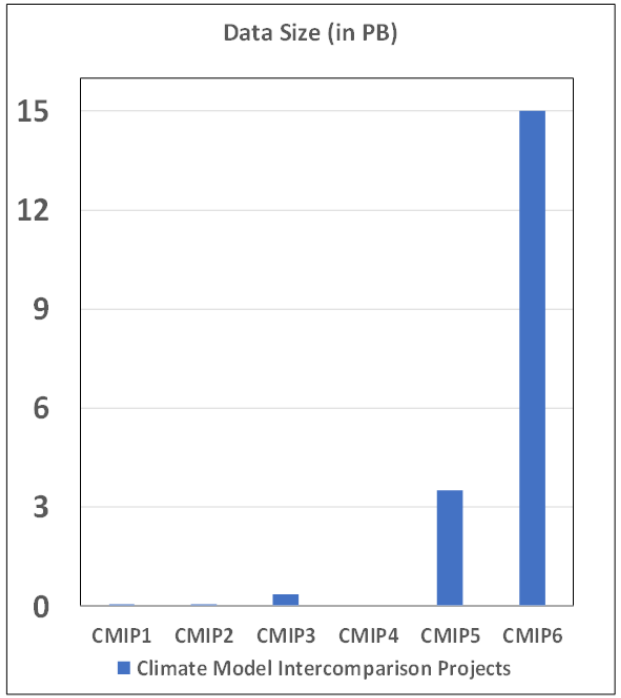
Recommended: https://github.com/nci-training/notebooks-python

Objectives:

- Wrap up of last week
- Cloud Computing
- OGC - Standards
- Climate Services Information Systems

info@nilshempelmann.de

Growing Amount of Data (Example Climate Model data)



Climate Data	Data Size
AMIP	?
CMIP1	1 GB
CMIP2	500 GB
CMIP3	35 TB
CMIP4	Not existing
CMIP5	3.5 PB
CMIP6	10-20 PB

Report
IPCC AR1 1990
IPCC AR2 1995
IPCC AR3 2001
IPCC AR4 2007
IPCC AR5 2014
IPCC AR6 expected: June 2022

Donnes OT	Data Size
Modis	?

GLO first edition 2017

→ = optimized science delivery

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

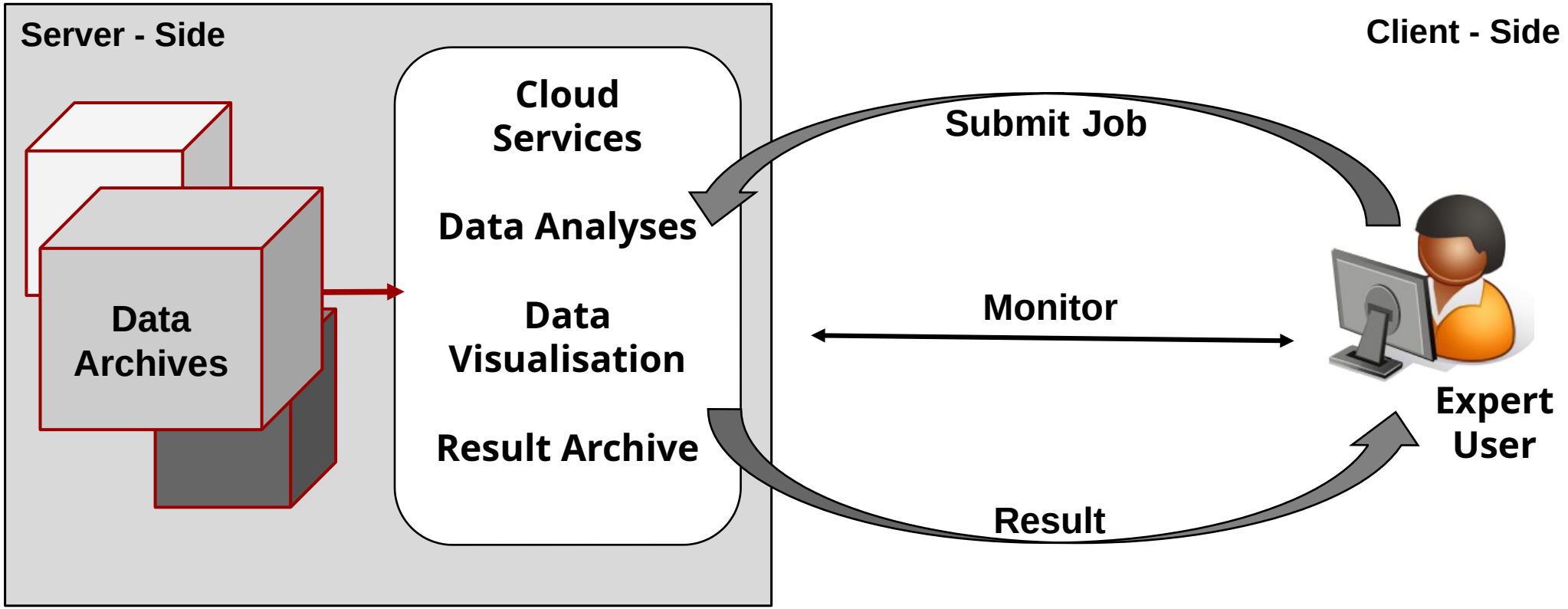
Internet connection (= Bandwidth)

info@nilshempelmann.de

Web Processing Service

high performance environment

(low internet bandwidth)

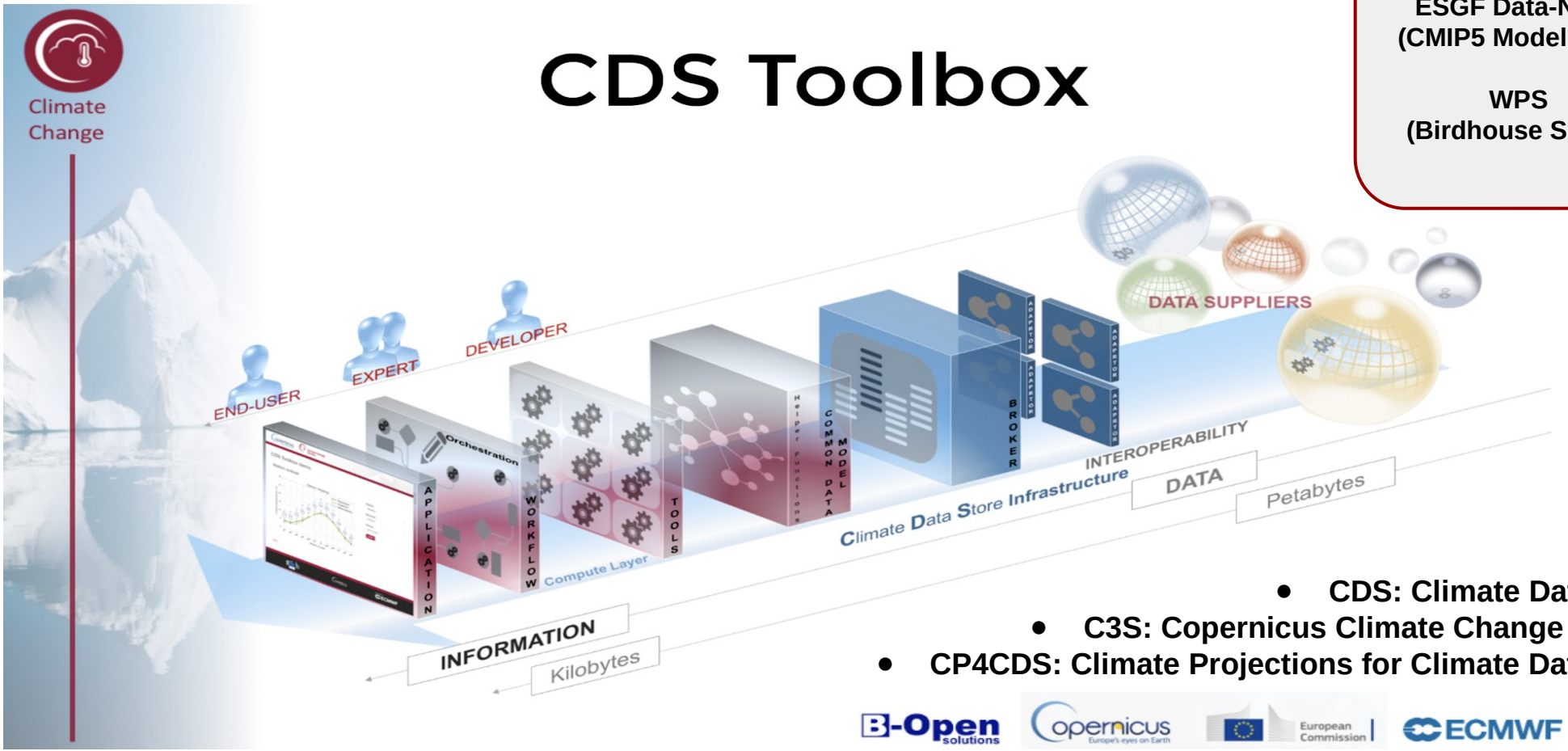


Copernicus Climate Change Service

CP4CDS

ESGF Data-Node (CMIP5 Model Data)

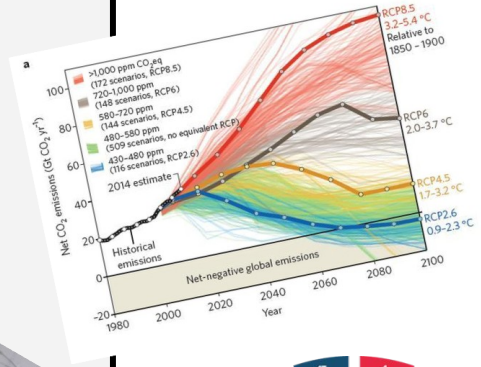
WPS (Birdhouse Setup)



- **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**

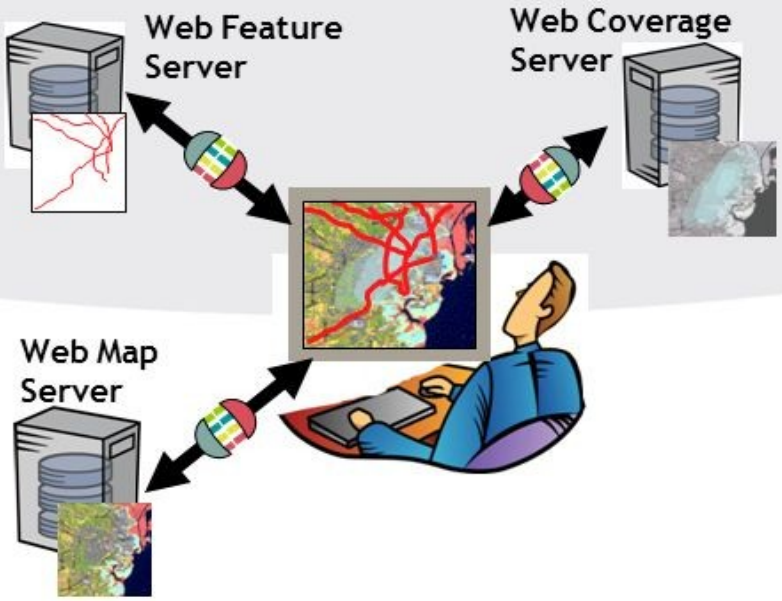


info@nilshempelmann.de

OGC Standard

Recommended by the UNGGIM !

The geospatial web is enabled by OGC standards:



- 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

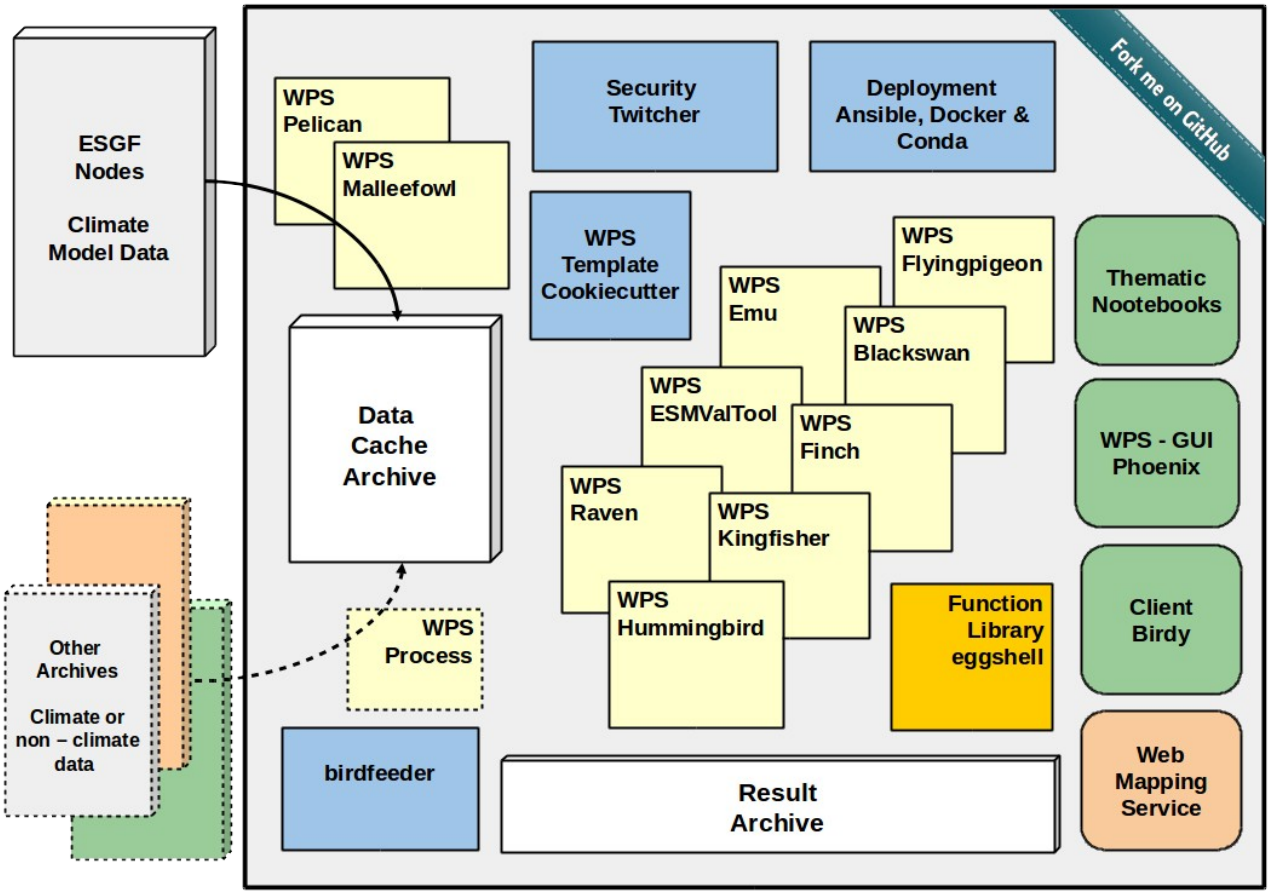
Osservare per prevedere, prevedere per prevenire

info@nilshempelmann.de

Birdhouse Ecosystem

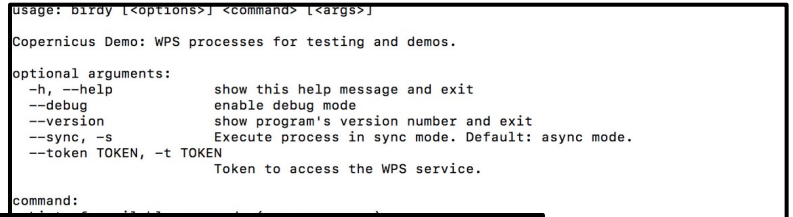
https://medium.com/birdhouse-newsletter

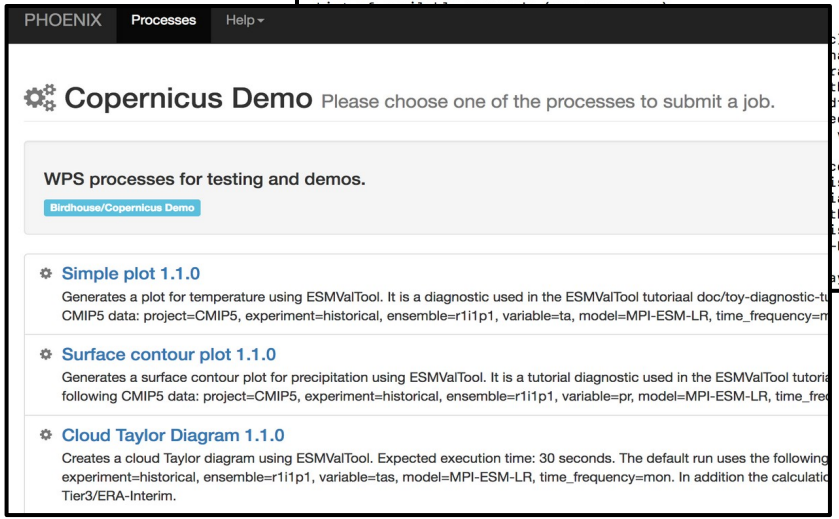
Open Source Software compartements to set up Open Standared based Web Processing Services (WPS)

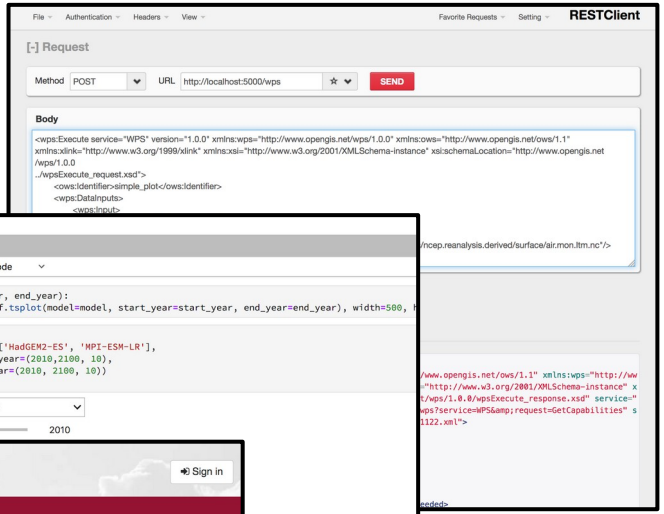


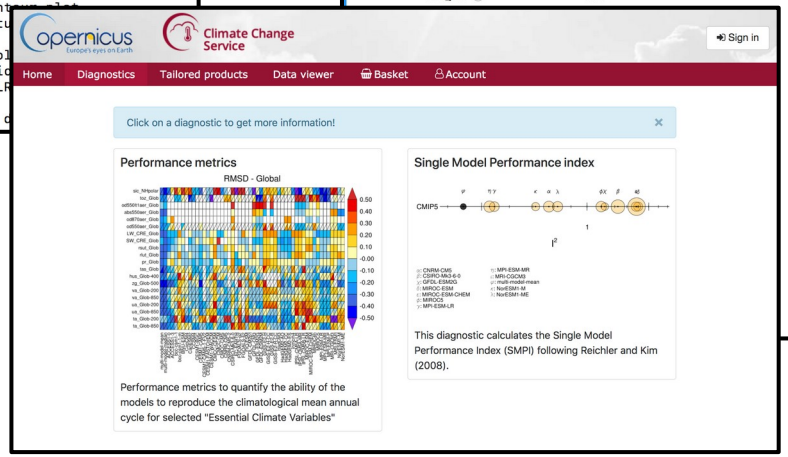
info@nilshempelmann.de

Client Side: Many Views on WPS



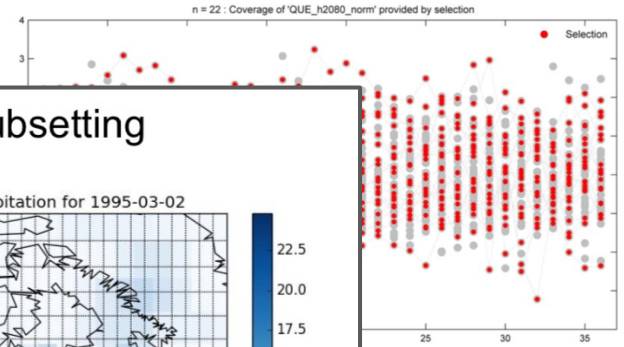




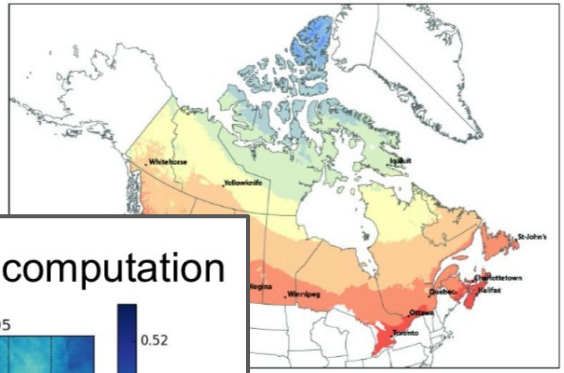


Example of Processes & Workflows

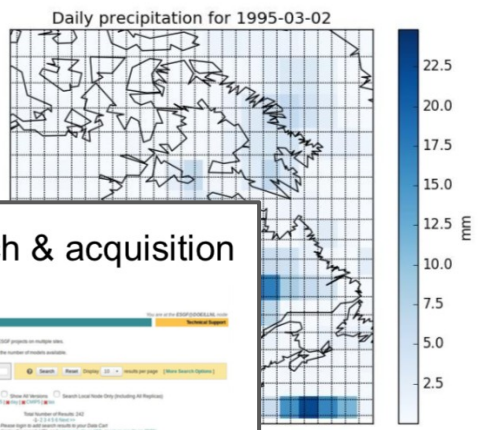
Selection of an ensemble of simulations



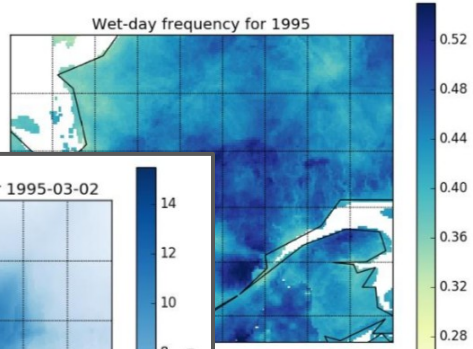
User specific visualization



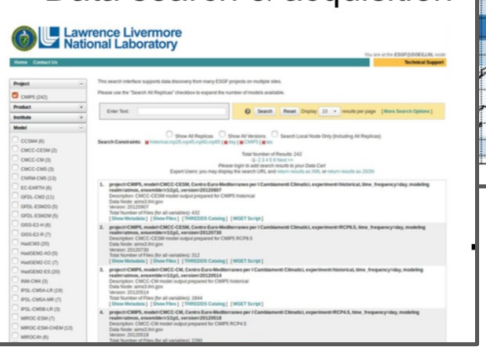
Subsetting

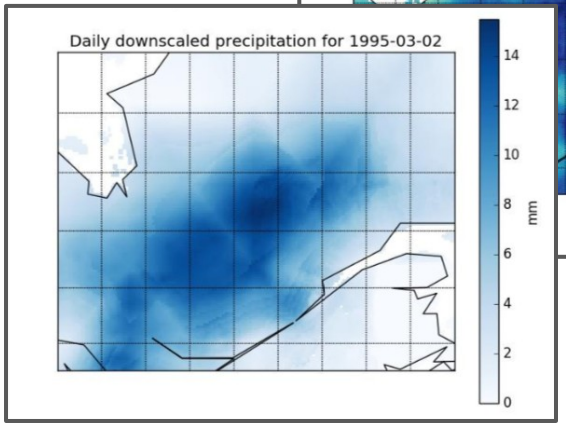


Climate indicator computation



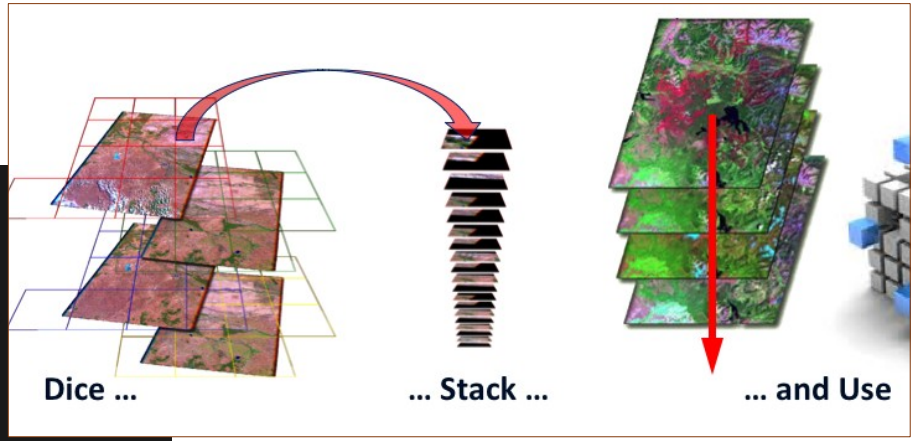
Data search & acquisition





info@nilshempelmann.de

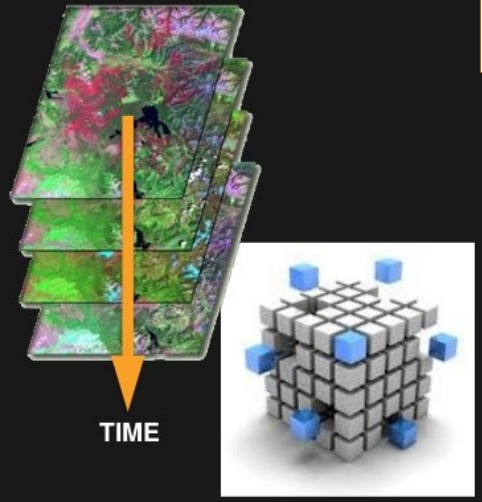
Open Data Cube



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



https://github.com/opendatacube

Including example notebooks

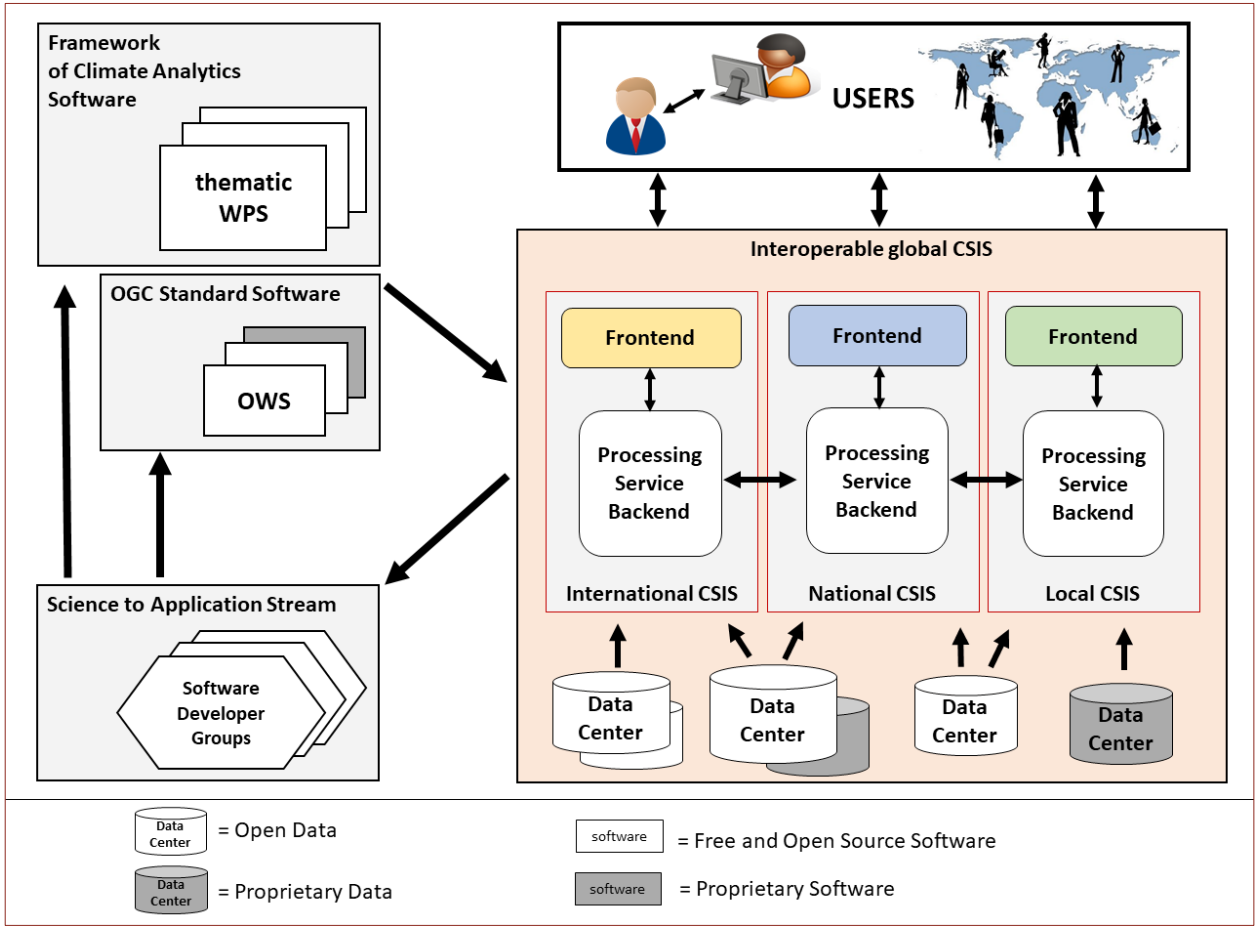
info@nilshempelmann.de

Open Data Cube Initiative

https://www.opendatacube.org/

info@nilshempelmann.de

Climate Services Information System



info@nilshempelmann.de

Hands ON:

Design a national CSIS Climate indice with WPS

info@nilshempelmann.de