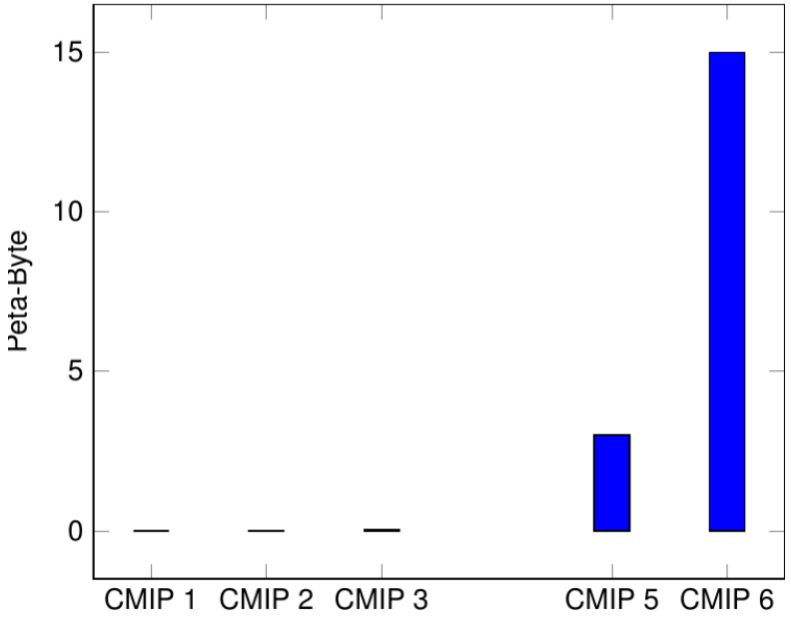
Projects based on the Web Processing Service framework birdhouse

Carsten Ehbrecht[1], Tom Landry[2], Nils Hempelmann[2], David Huard[3], Stephan Kindermann[1]

1. German Climate Computing Center (DKRZ), Germany
2. Computer Research Institute of Montreal (CRIM), Canada
3. Ouranos, Canada

Growing Amount of Data (Example Climate Model data)



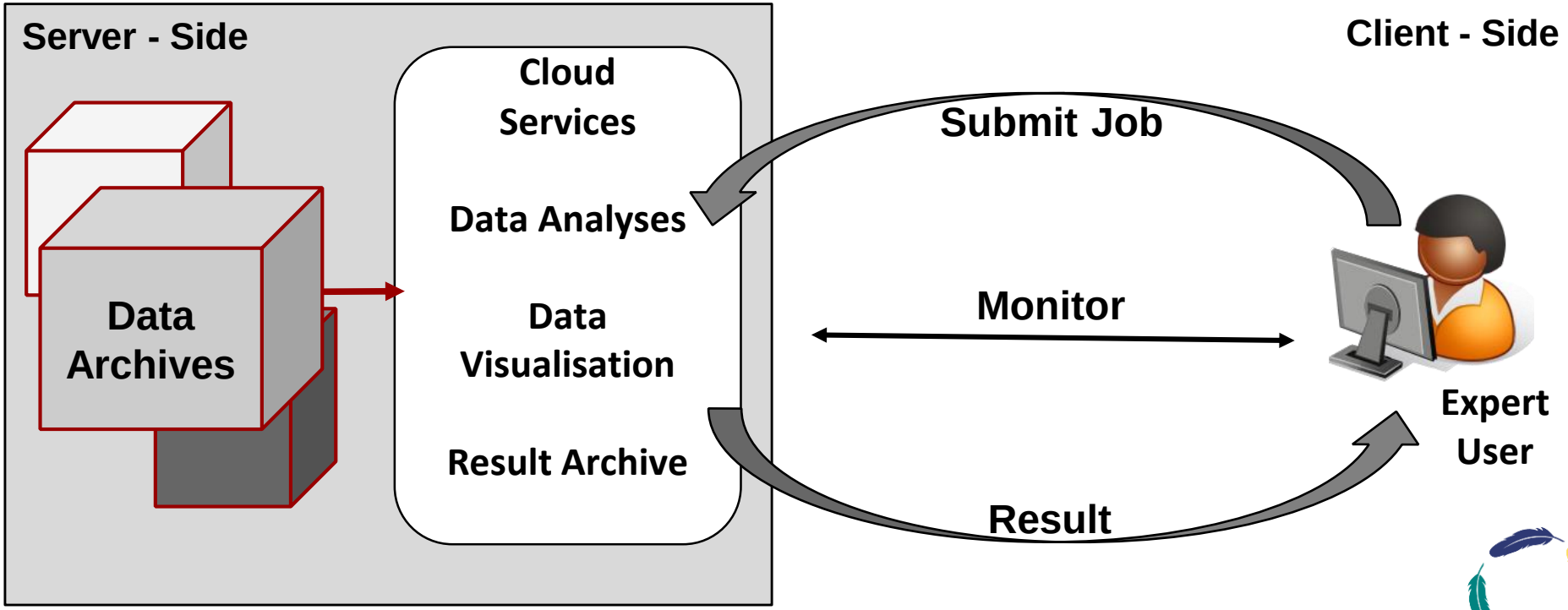
CMIP 1	1 GB	IPCC AR 1 1990
CMIP 2	500 GB	IPCC AR 2 1995
CMIP 3	35 TB	IPCC AR 3 2001
	Not existing	IPCC AR 4 2007
CMIP 5	3,5 PB	IPCC AR 5 2014
CMIP 6	10-20 PB (in ESGF)	IPCC AR 6 12-16 April 2021

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

Web Processing Service

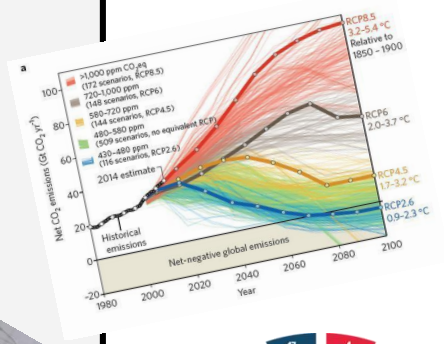
high performance environment

(low internet bandwidth)



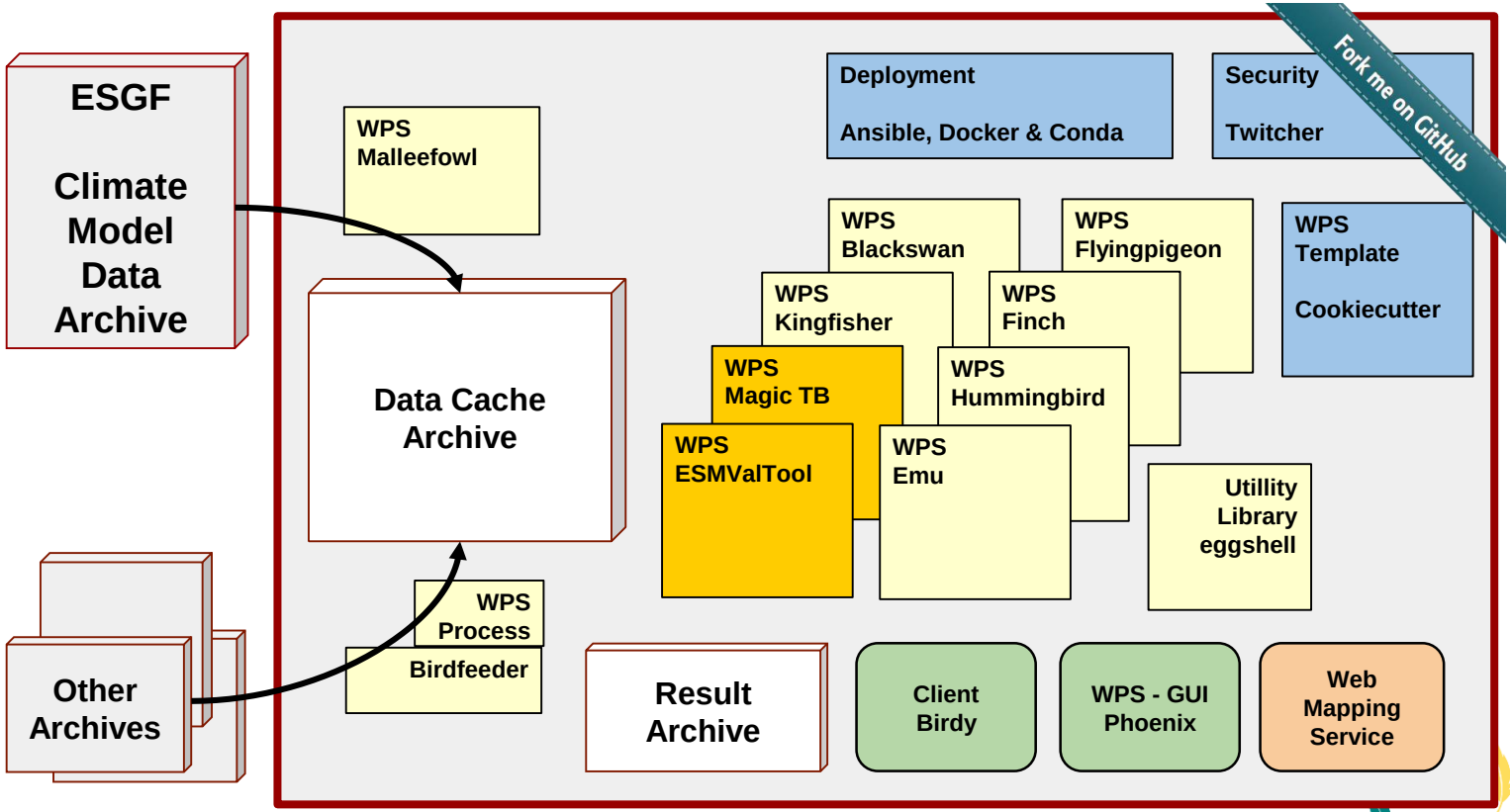
Advantage of Server-Side Services

- **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**

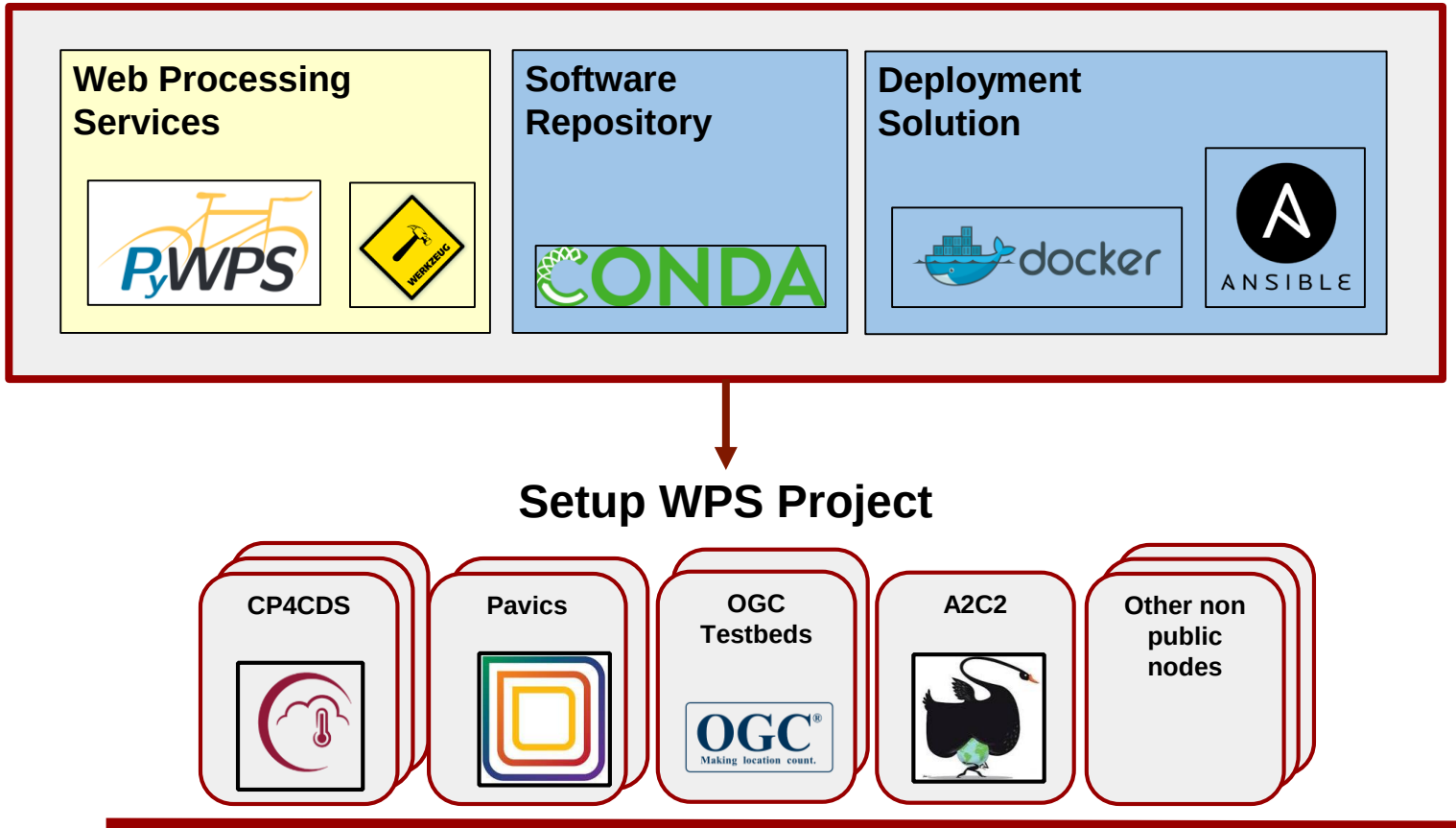


Birdhouse Ecosystem

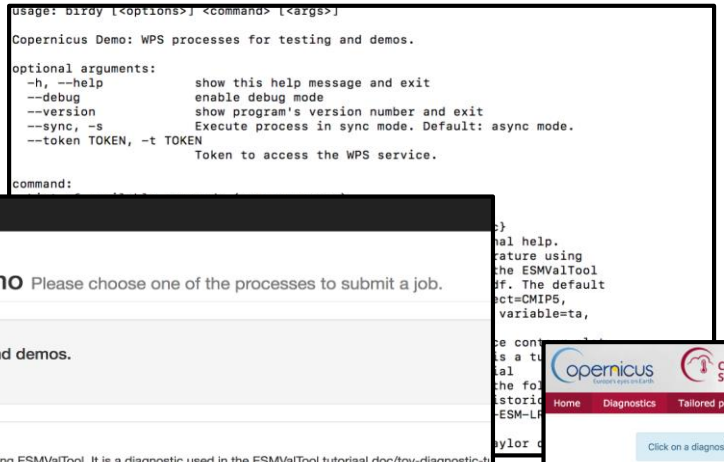
EMU
Test suite
Flyingpigeon
Sandbox
Malleefowl
ESGF search&fetch
Blackswan
Extreme Event assessment
Kingfisher
EO-Data
Hummingbird
Data Quality Control
Finch
Climate Impact
ESMValTool
EarthSystem Validation
Magic Toolbox
Climate indices

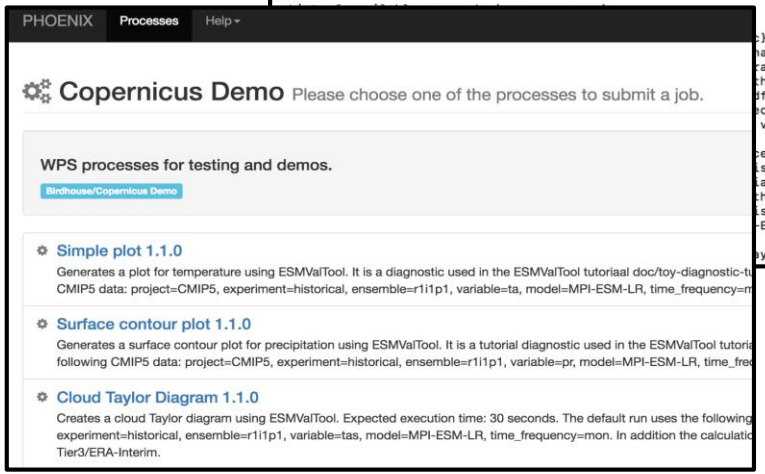


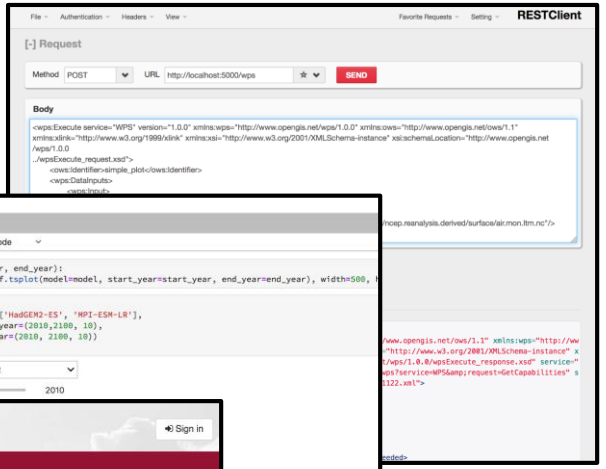
Server Side: Deployment

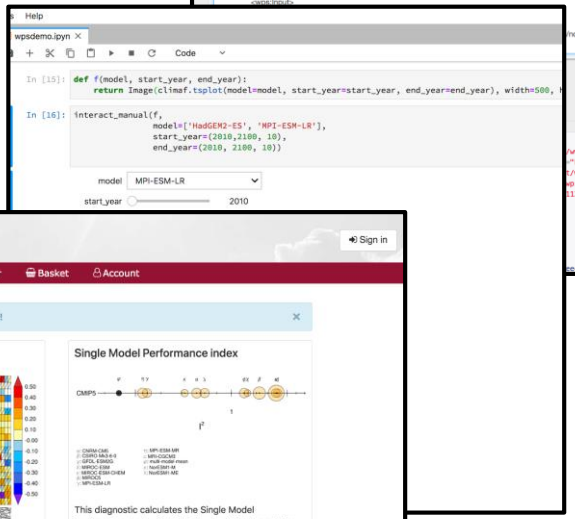


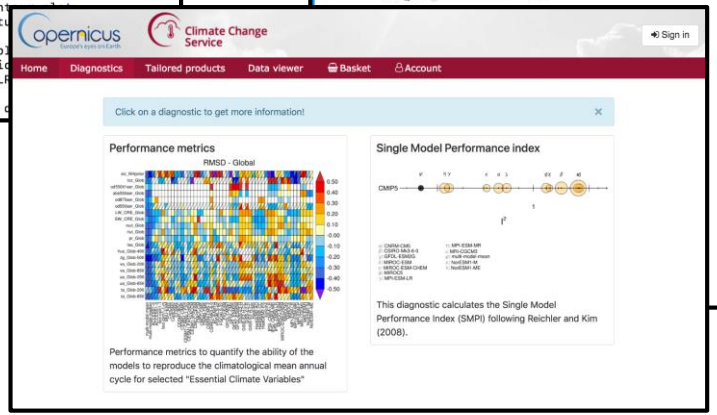
Client Side: Many Views on WPS



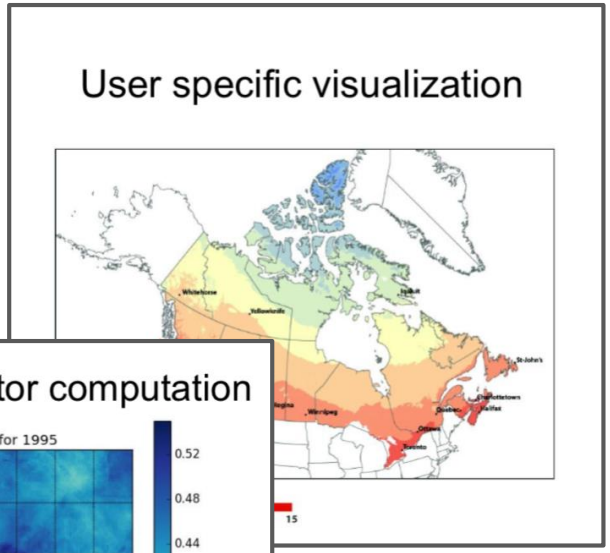


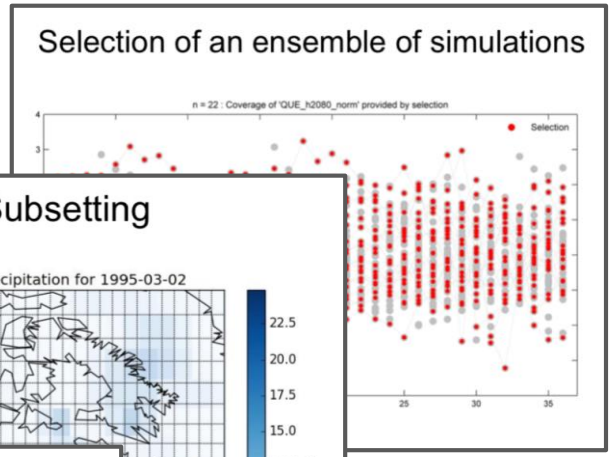


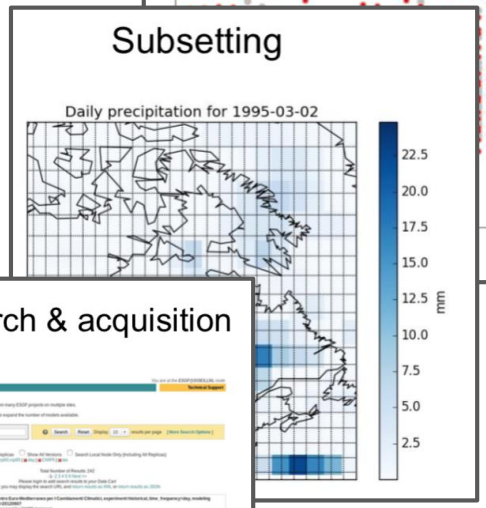


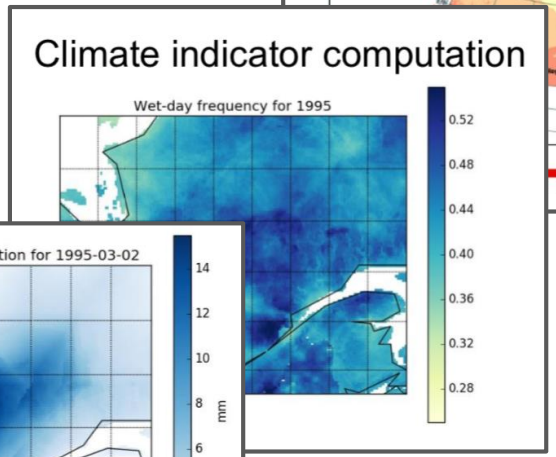


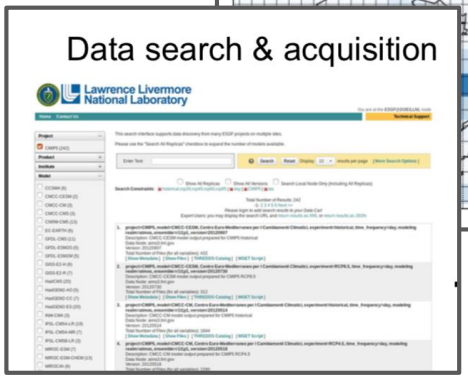
Example of Processes & Workflows

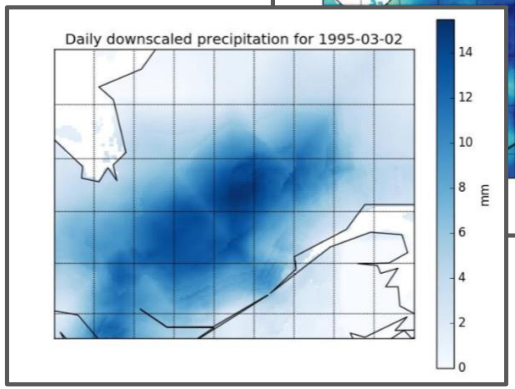




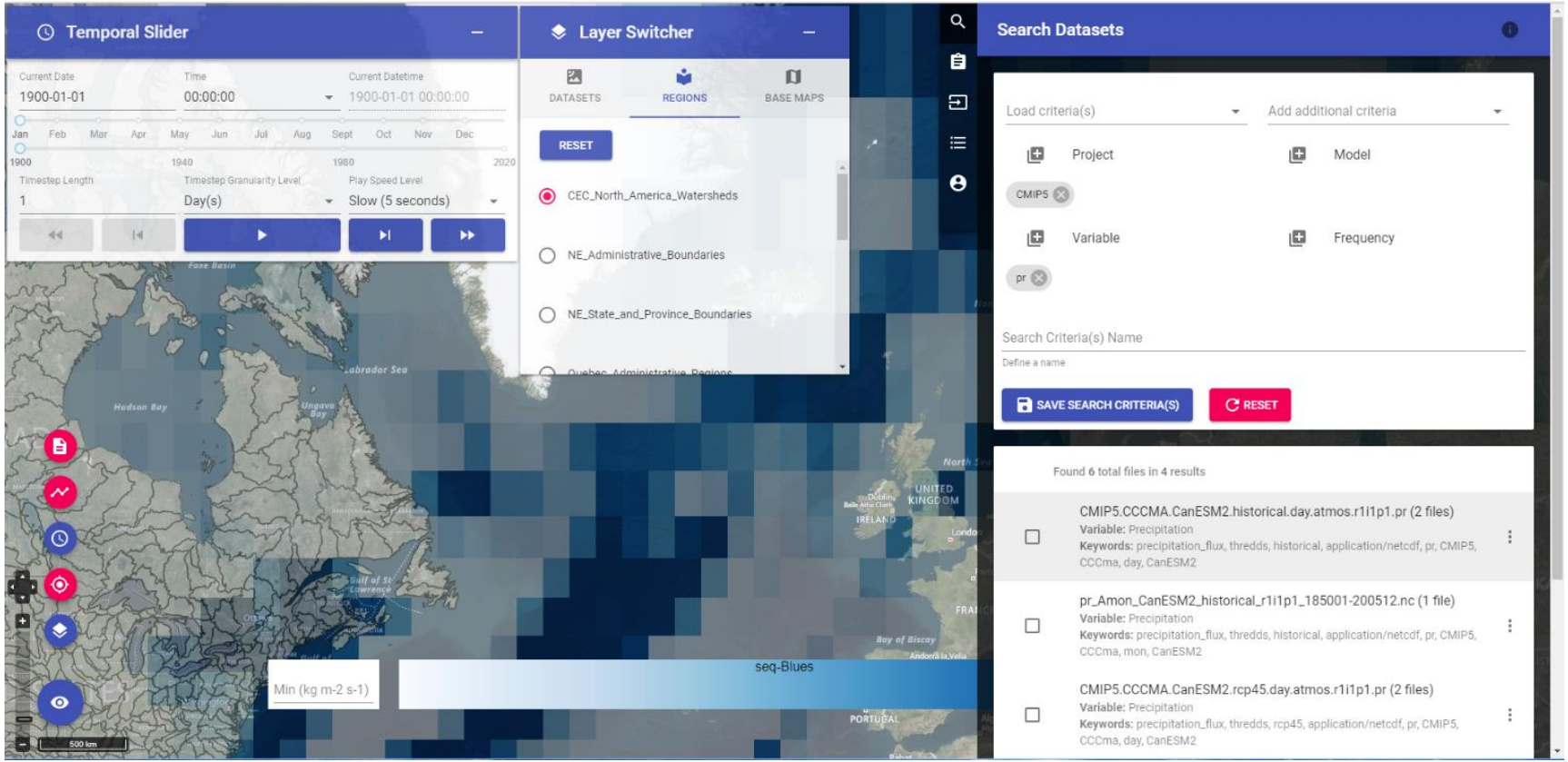








Project: PAVICS



Project: PAVICS NexGen

- Canadian-led Science Gateway
- Will provide a federated cyberinfrastructure
- Advanced hydro & climate services, Machine Learning and Earth Observation tools
- Geospatial Interoperability through OGC Web Services

PAVICS

Information
- Research platform
- Climate portal
- Sectorial modules

Services
- Datacube
- Machine Learning
- Climate services
- Hydrological services

Middleware
- Load balanced proxy
- Authentication
- Authorisation

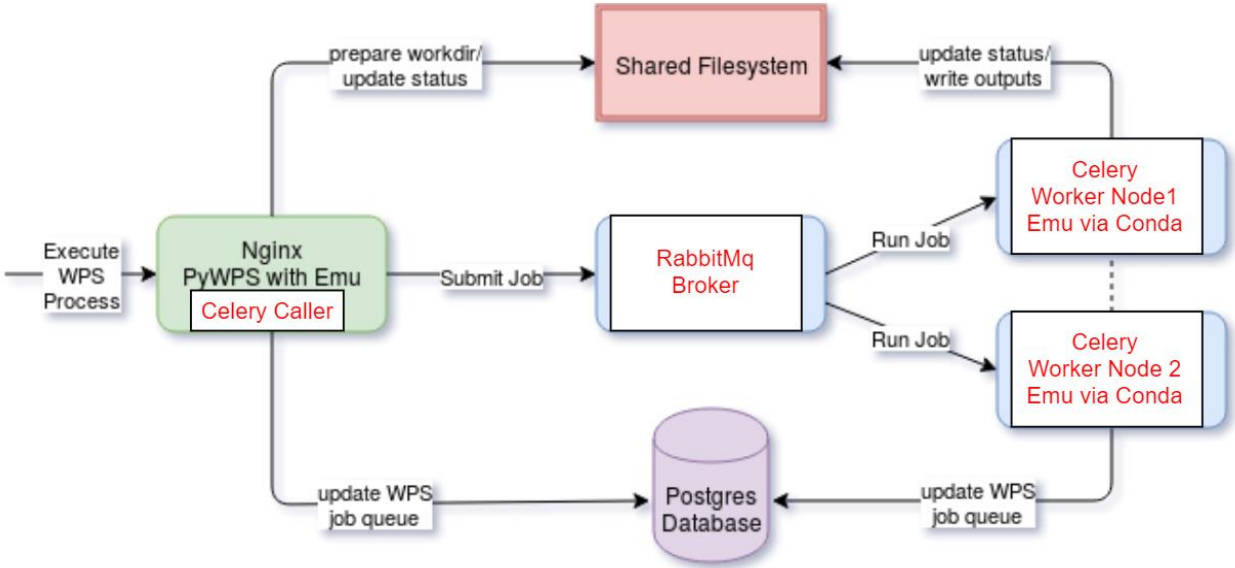
Interoperability
- OGC Web Services (OWS)

Computation
- Hybrid cloud
- HPC

Data
- Climate data
- Earth Observation data

Project: OGC Testbeds

- **OGC Testbeds:**
  - Collaborative efforts to define, design, develop, and test candidate interface specifications.
- **Testbed-13:**
  - Builds on Birdhouse and PyWPS HPC scheduler
  - Implements cloud computing & application packaging (Docker)
- **Testbed-14:**
  - Extends Birdhouse security and job processing (OpenAPI)
  - Implements interoperable workflows (CWL, OWS Context)



Project: Copernicus Climate

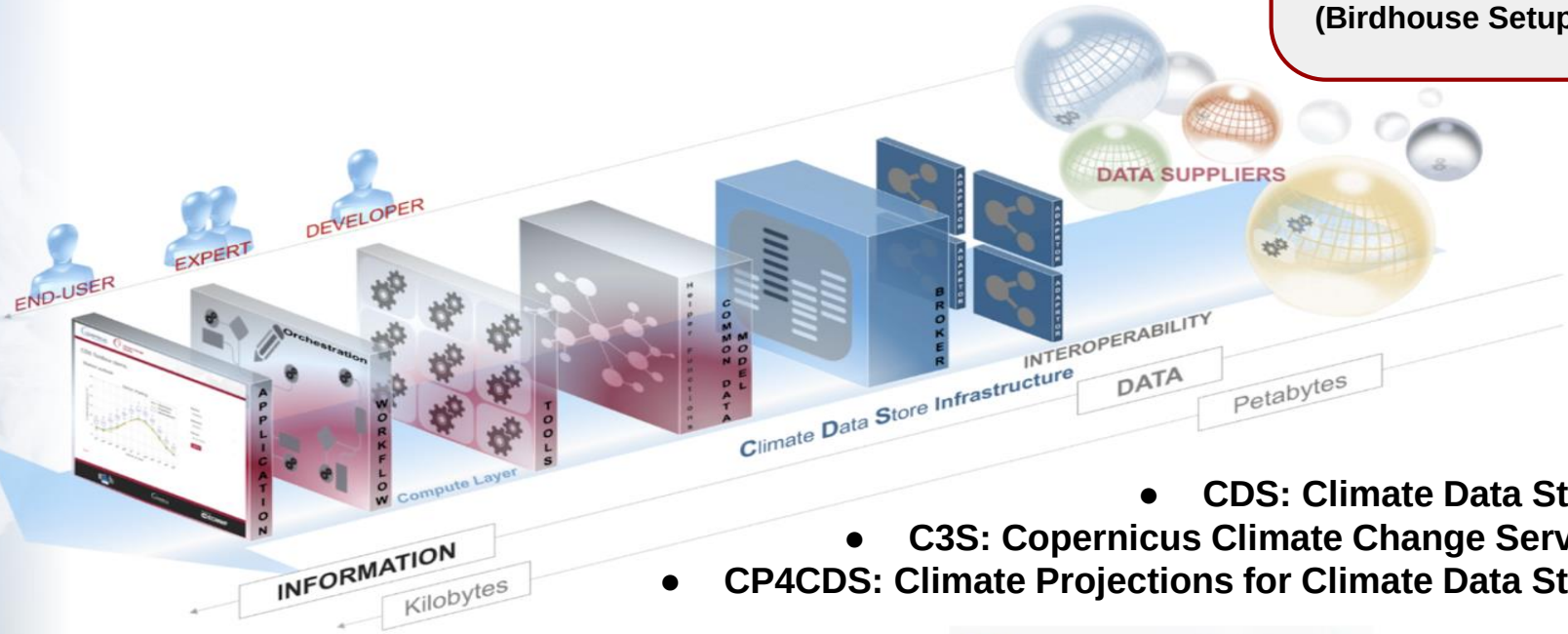
Climate Change

CP4CDS

ESGF Data-Node (CMIP5 Model Data)

WPS (Birdhouse Setup)

CDS Toolbox

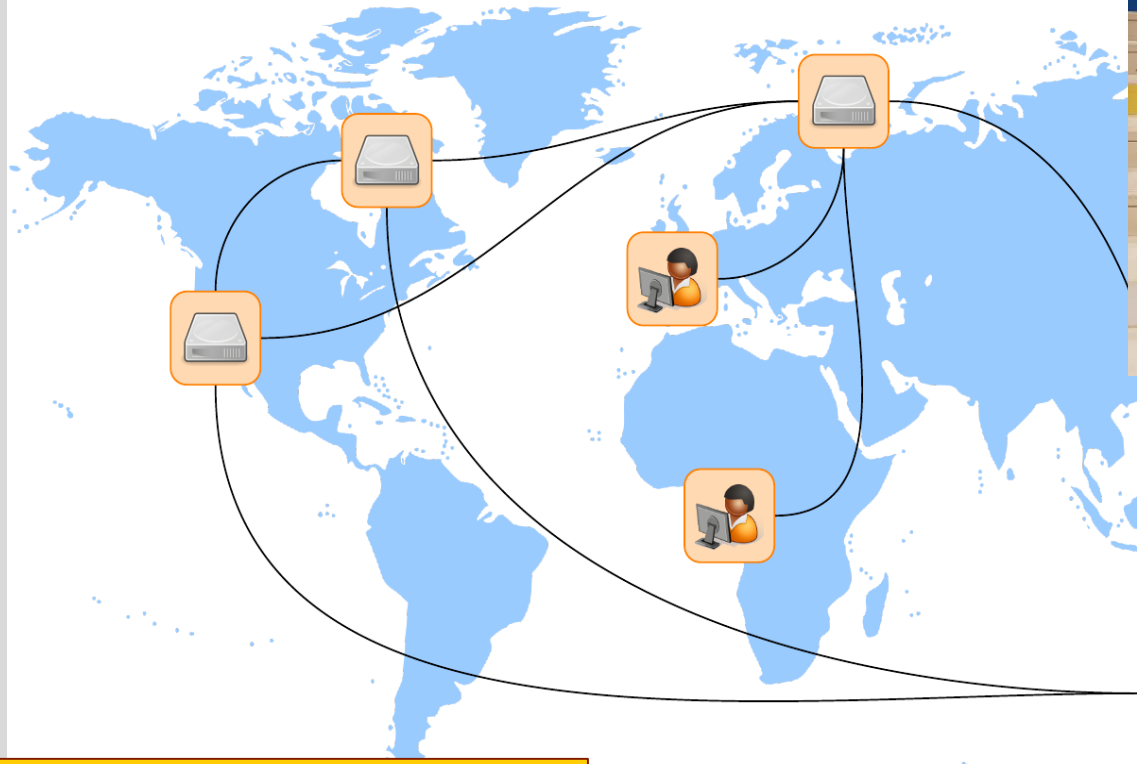


- CDS: Climate Data Store
- C3S: Copernicus Climate Change Service
- CP4CDS: Climate Projections for Climate Data Store

B-Open solutions | Copernicus Europe's eyes on Earth | European Commission | ECMWF

Harnessing WPS for Sustainable Development

- Indicators for the Sustainable development Goals (SDGs)
- Monitoring and Reporting required from UNFCCC, UNCCD and CBD
- non binding Sendai Framework for Disaster Risk Reduction - UNISDR
- etc ...



Leave no one behind

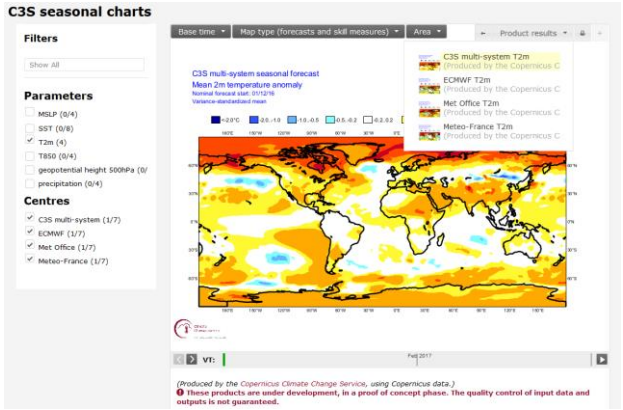
Outlook

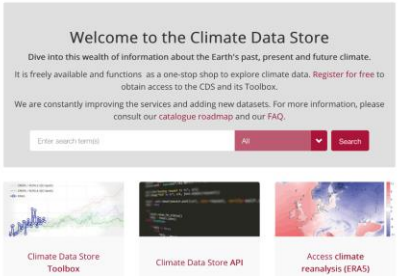
- **EU COPERNICUS**
  - C3S with CP4CDS ready for production use in 2019.
    (based on CMIP5 data)
  - extended for CORDEX (regional model data) in 2019/2020
  - extended for CMIP6 … ?
- **Canada**
  - Govt of Canada: Canadian Center for Climate Services (CCCS)
  - Pan-Canadian federated cyberinfrastructure
  - Earth Observation support, integration of ML/DL
- WPS on **ESGF Nodes**
- **A2C2** Service for Extreme Weather Assessment
- …. etc.

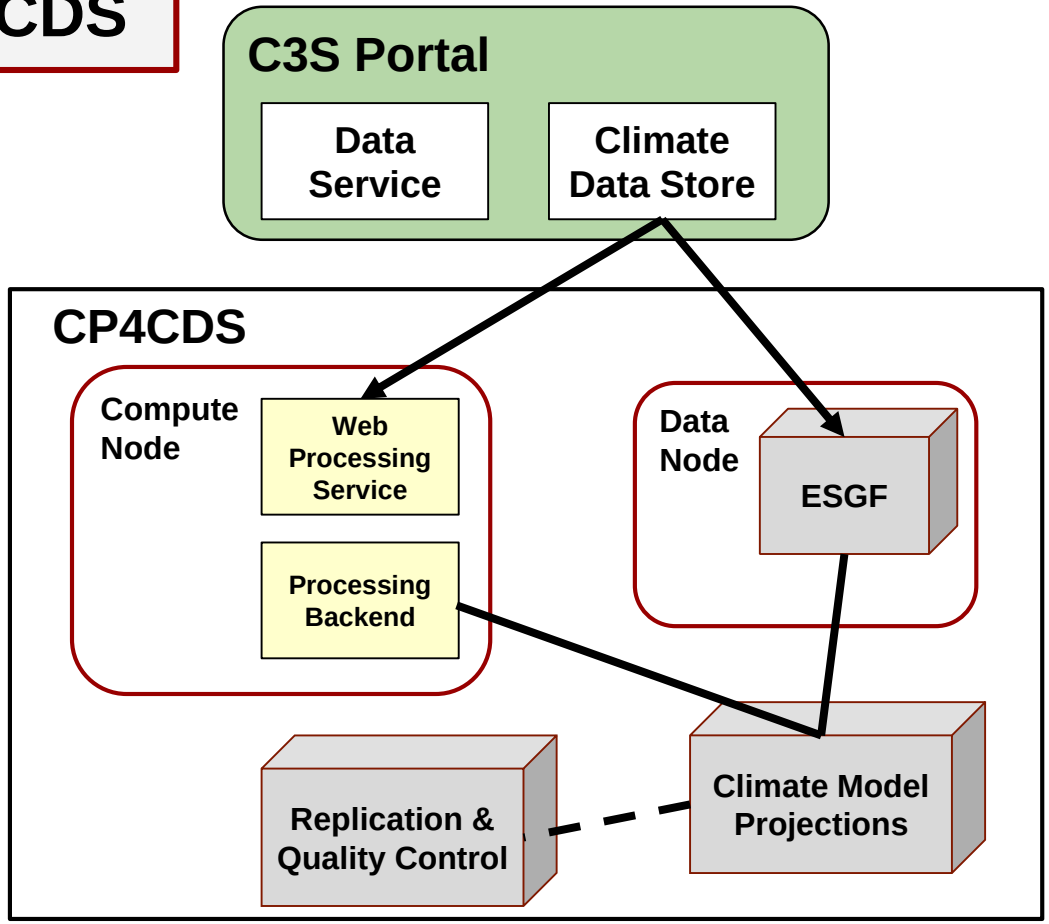
→ http://bird-house.github.io/

Deleted slides

Projects: C3S and CP4CDS



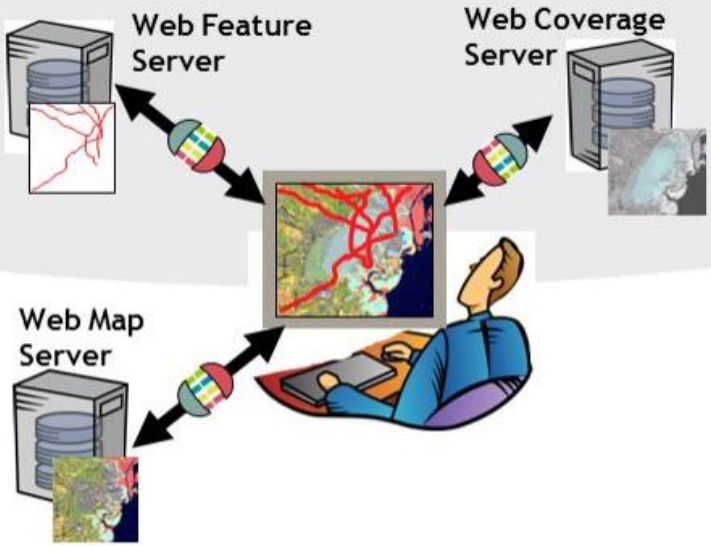




Open Geospatial Consortium

OGC®
Making location count.

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…

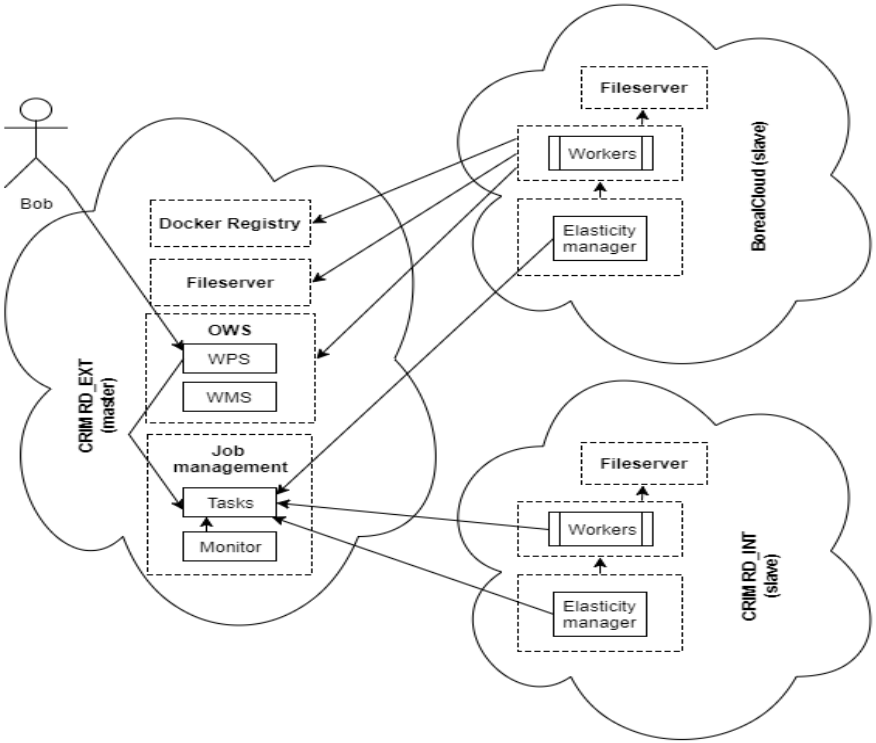
OURANOS

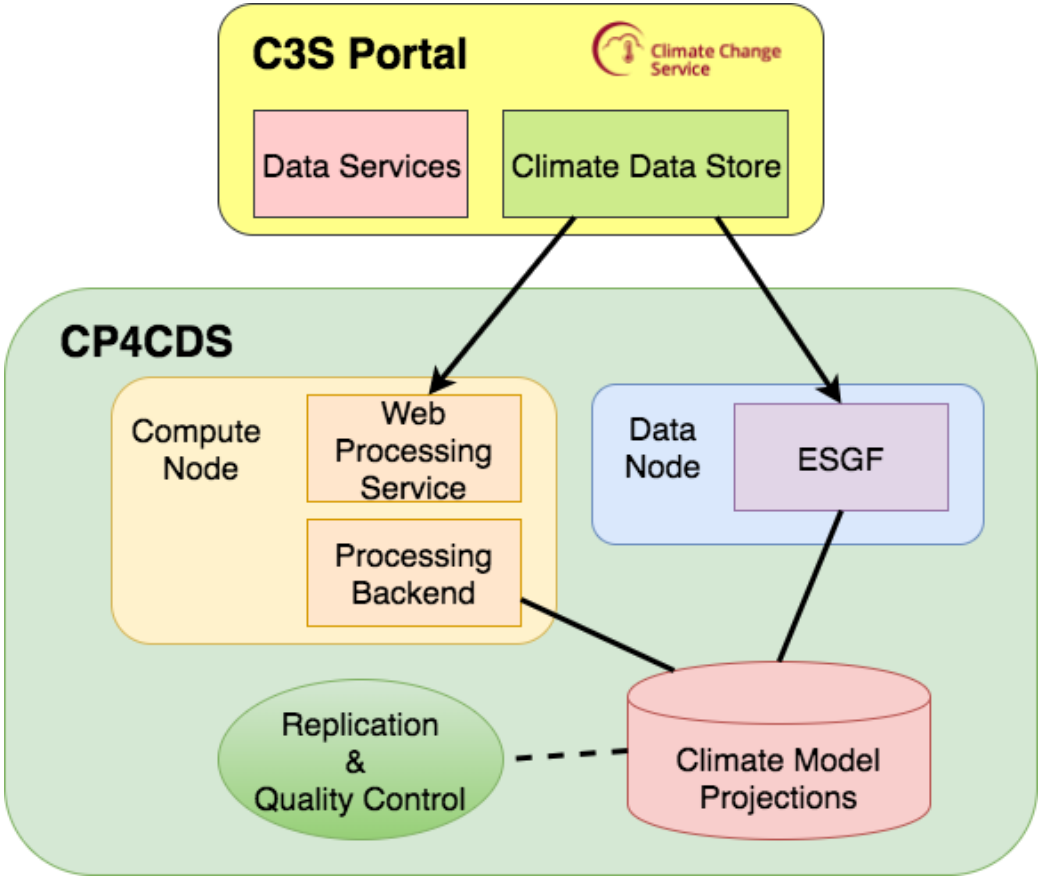
DKRZ
DEUTSCHES KLIMARECHENZENTRUM

CRIM

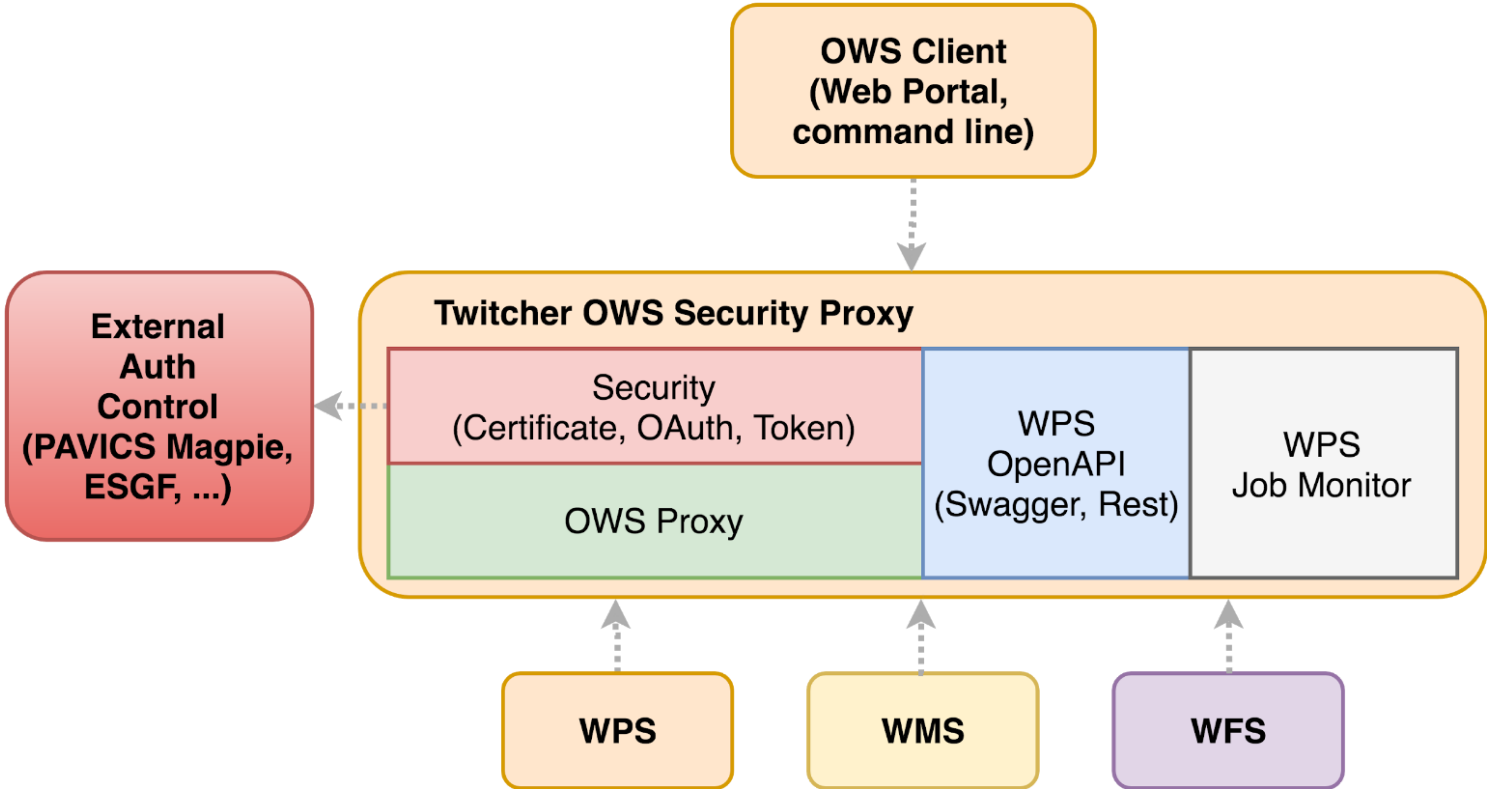
Relevant to geospatial applications: Critical Infrastructure, Emergency Management, Weather, Climate, Homeland Security, Defense & Intelligence, Oceans Science, etc

Osservare per prevedere, prevedere per prevenire


Bob
Docker Registry
Fileserver
OWS
WPS
WMS
Job management
Tasks
Monitor
CRIM RD_EXT (master)
Fileserver
Workers
Elasticity manager
BorealCloud (slave)
Fileserver
Workers
Elasticity manager
CRIM RD_INT (slave)


C3S Portal
Climate Change Service
Data Services
Climate Data Store
CP4CDS
Compute Node
Web Processing Service
Processing Backend
Data Node
ESGF
Replication & Quality Control
Climate Model Projections

Security: Twitcher OWS Proxy



Earth System Grid Federation (ESGF)

ESGF
Earth System Grid Federation

R&D Area	ESG-I 1999-2001	ESG-II 2001-2006	ESG-CET 2006-2011	ESG-F 2011-2020	ESG-VL 2020-
Data Management					
User Interface and Search					
Hardware & Network					
Data Transfer					
Installation (Containerized)					
Authentication & Authorization					
Federation					
Version Control					
Quality Control & Assurance					
Replication					
Distributed Search					
Metrics					
User Notification					
Long-tail Publication					
Distributed Computation					
Data Citation					
Provenance Capture					
Workflow					
Dynamic Resources					
In situ Analysis					
Machine Learning					
Uncertainty Quantification					
Cloud Development					
Analytical Modeling					
Knowledge Discovery					

Petabytes (10^{15}) Model Intercomparison Projects

Planned ESGF Data Archival Capabilities

Exabytes (10^{15}) Examples Model Simulations, *In Situ*, Remotes Sensing Observations, and Reanalysis

Climatology, Diagnostics, Ecosystem, Hydrology, Biology, Forestry.

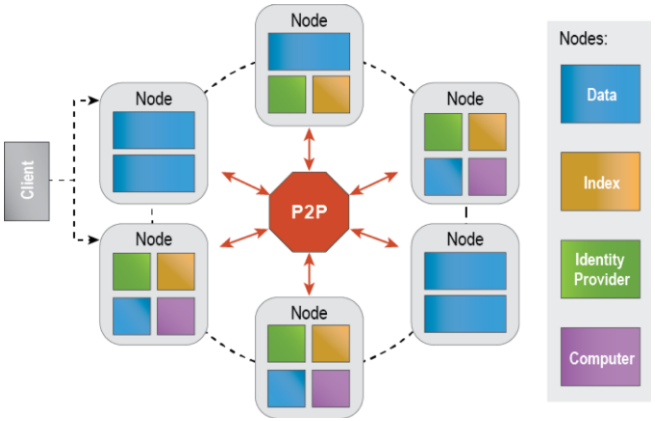
Usable Capability | Prototype Capability | Research

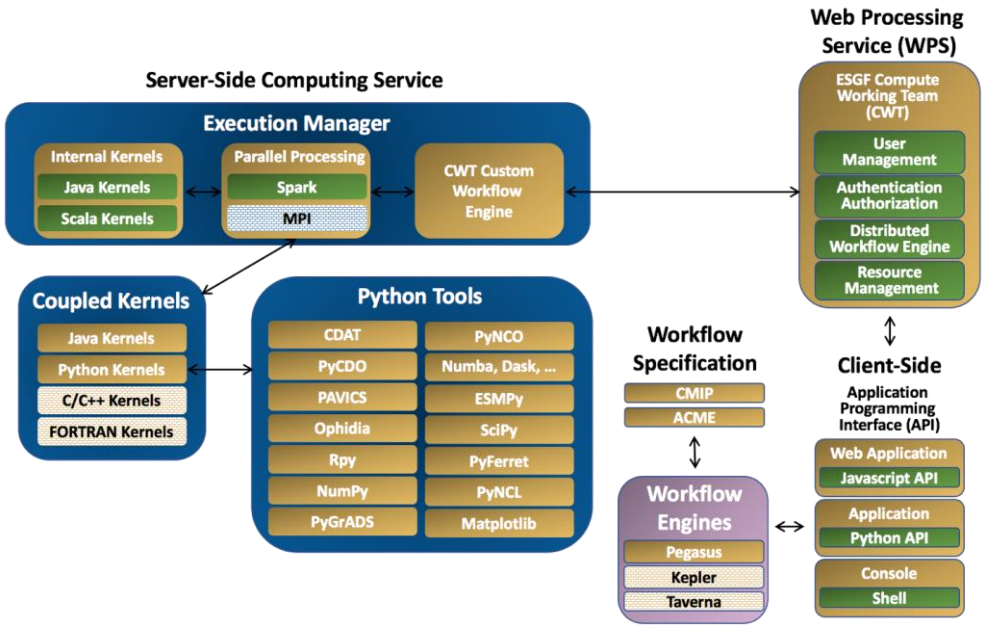
Downloads by Continent and Country

https://esgf.llnl.gov

Project: Connectivity to ESGF

ESGF
Earth System Grid Federation



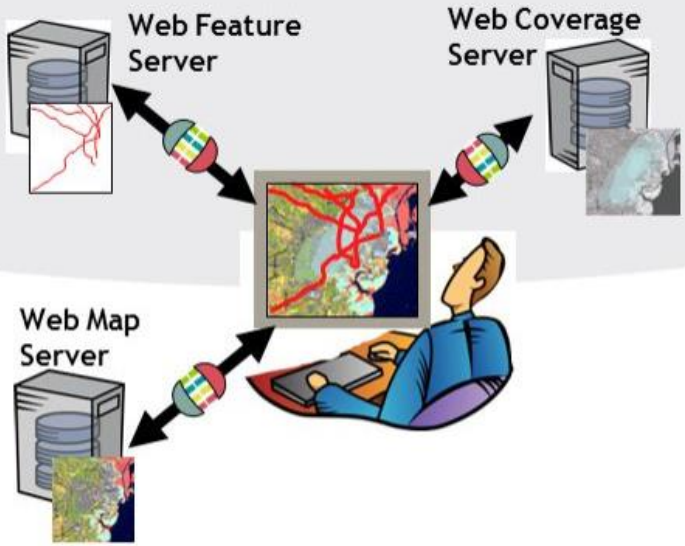


https://esgf.llnl.gov

Open Geospatial Consortium

OGC®
Making location count.

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…

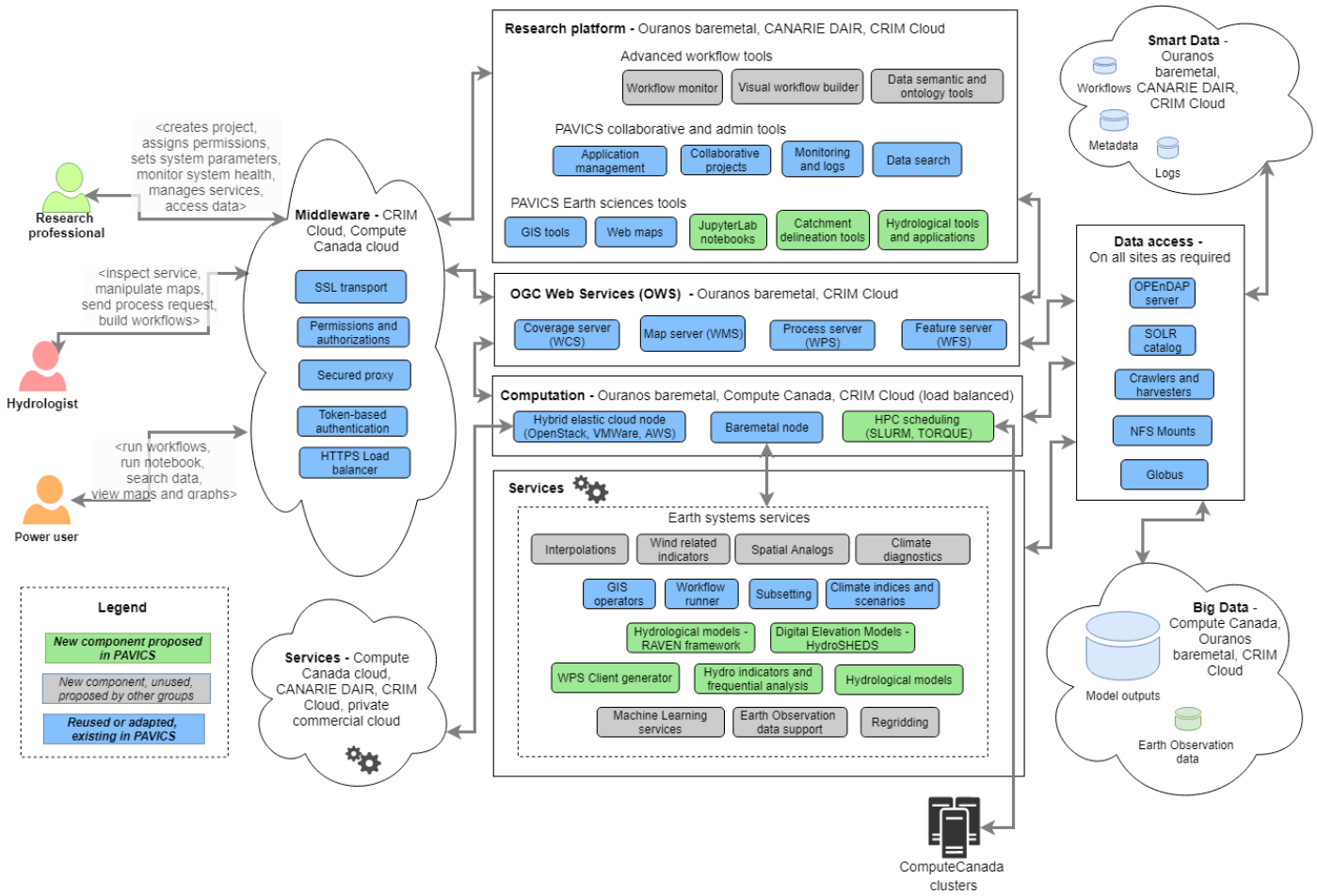
Relevant to geospatial applications: Critical Infrastructure, Emergency Management, Weather, Climate, Homeland Security, Defense & Intelligence, Oceans Science, etc

Osservare per prevedere, prevedere per prevenire

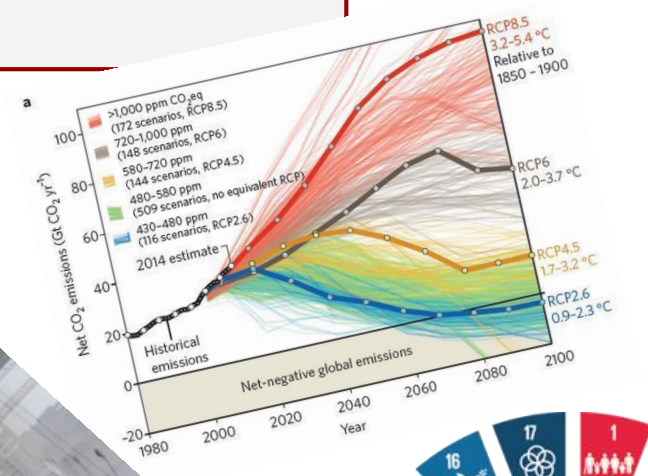
OURANOS

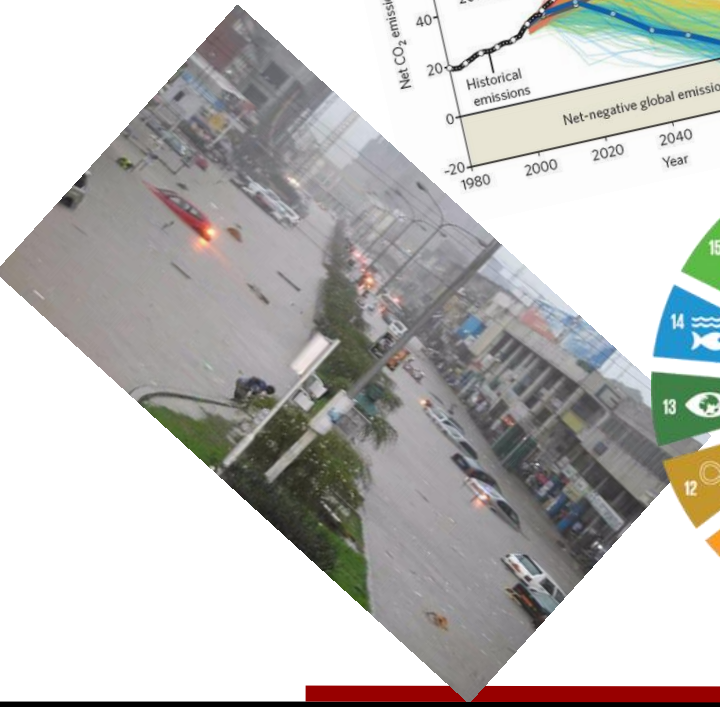
DKRZ
DEUTSCHES KLIMARECHENZENTRUM

CRIM


Research platform - Ouranos baremetal, CANARIE DAIR, CRIM Cloud
Advanced workflow tools
Workflow monitor
Visual workflow builder
Data semantic and ontology tools
PAVICS collaborative and admin tools
Application management
Collaborative projects
Monitoring and logs
Data search
PAVICS Earth sciences tools
GIS tools
Web maps
JupyterLab notebooks
Catchment delineation tools
Hydrological tools and applications
Smart Data - Ouranos baremetal, CANARIE DAIR, CRIM Cloud
Workflows
Metadata
Logs
<creates project, assigns permissions, sets system parameters, monitor system health, manages services, access data>
Research professional
Middleware - CRIM Cloud, Compute Canada cloud
SSL transport
Permissions and authorizations
Secured proxy
Token-based authentication
HTTPS Load balancer
<inspect service, manipulate maps, send process request, build workflows>
Hydrologist
<run workflows, run notebook, search data, view maps and graphs>
Power user
OGC Web Services (OWS) - Ouranos baremetal, CRIM Cloud
Coverage server (WCS)
Map server (WMS)
Process server (WPS)
Feature server (WFS)
Computation - Ouranos baremetal, Compute Canada, CRIM Cloud (load balanced)
Hybrid elastic cloud node (OpenStack, VMWare, AWS)
Baremetal node
HPC scheduling (SLURM, TORQUE)
Services
Earth systems services
Interpolations
Wind related indicators
Spatial Analogs
Climate diagnostics
GIS operators
Workflow runner
Subsetting
Climate indices and scenarios
Hydrological models - RAVEN framework
Digital Elevation Models - HydroSHEDS
WPS Client generator
Hydro indicators and frequential analysis
Hydrological models
Machine Learning services
Earth Observation data support
Regridding
Data access - On all sites as required
OPENDAP server
SOLR catalog
Crawlers and harvesters
NFS Mounts
Globus
Big Data - Compute Canada, Ouranos baremetal, CRIM Cloud
Model outputs
Earth Observation data
Legend
New component proposed in PAVICS
New component, unused, proposed by other groups
Reused or adapted, existing in PAVICS
Services - Compute Canada cloud, CANARIE DAIR, CRIM Cloud, private commercial cloud
ComputeCanada clusters

Outline :





- Introduction
  - Data growth for SDG
  - Big Data Problem
  - Main Concept
  - OGC
- Birdhouse framework
  - Server-Client Side
  - Deployment
  - Security
  - Workflow
- Projects
  - C3S Portal
  - Pavics
  - A2C2
  - ESGF deployment
  - Earth observation
- Global picture Data4SDG
- Outlook