is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

European Geosciences Union 2015
Vienna, Austria

CAPACITIES

INTRO

Real use of open standards and technologies http://climate4impact.eu/

How open source software and open standards for data access and exploration are successfully applied in the "Climate for impact" portal.

CERFACS, **KNMI**, University of Cantabria, SMHI, Wageningen University, CMCC, STFC, IPSL

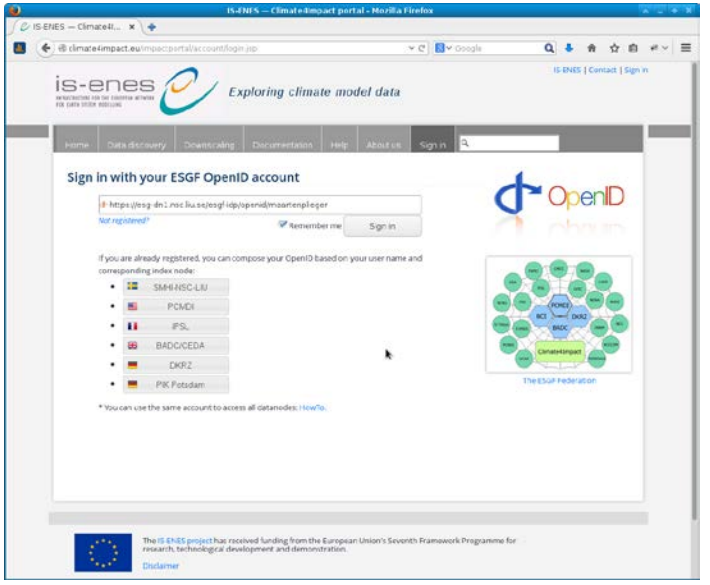
Maarten Plieger, Wim Som de Cerff, Ernst de Vreede,
Christian Pagé, Natalia Tatarinova,
Antonio Cofiño, Manuel Vega Saldarriaga,
Ronald Hutjes, Fokke de Jong,
Lars Bärring, Elin Sjokvist

CC BY

INTRO

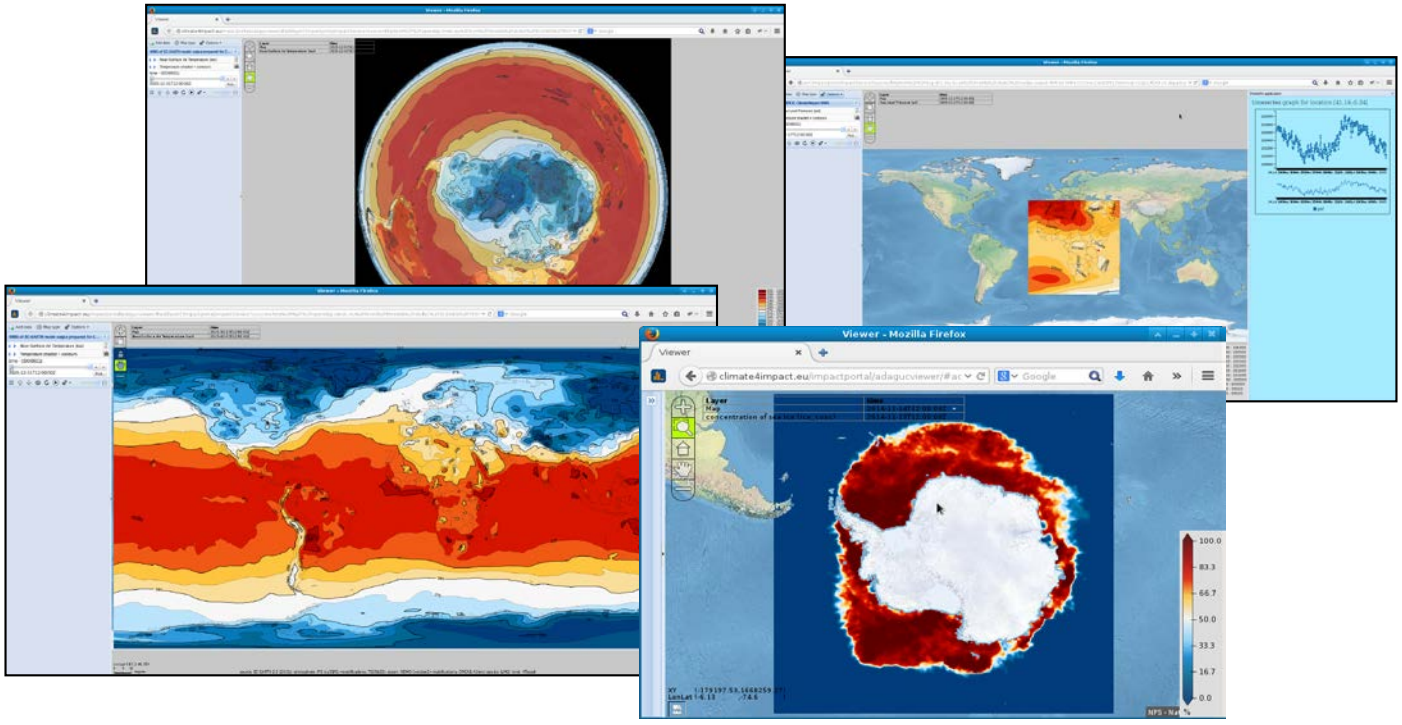
http://climate4impact.eu/

- Platform for impact researchers to explore climate data and perform analysis
- Current phase:
  - Implementing real use cases from impact researchers
  - Bulk download, bulk processing, etc…
- Search ESGF infrastructure for global climate models and regional climate models (CMIP5 / CORDEX)
- Visualize ESGF data - using ADAGUC Web Map Services
- Perform calculations / process data - PyWPS and ICCLIM
  - Climate indices calculation and data reduction
  - Personal store for processing outcomes
- Upload and store your own data
  - Visualize your own data online!
  - Process your own data online!



Real use of open standards and open source software:

- Data access over OPeNDAP → THREDDS
- Online analysis using Web Processing Services → PyWPS and ICCLIM
- Online visualization using Web Map Services → ADAGUC WMS
- Single Sign On using OpenId, delegation using MyProxy X509



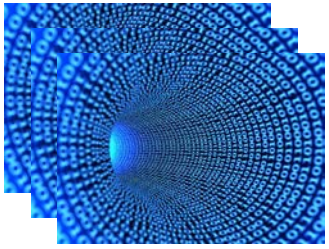
Interested? Come and take a look at PICO screen 1.2 !!!

CC BY

Data standards

Work done on client

Data files
HTTP / FTP

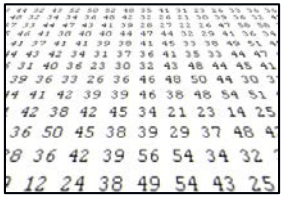


Array based data

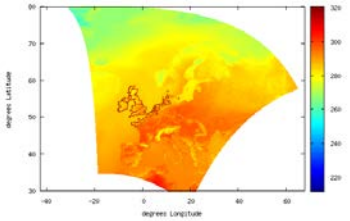
OPeNDAP

01001010010
01010010010
10010100100
10010011000

Referenced data
OGC WCS



Graphics
OGC WMS

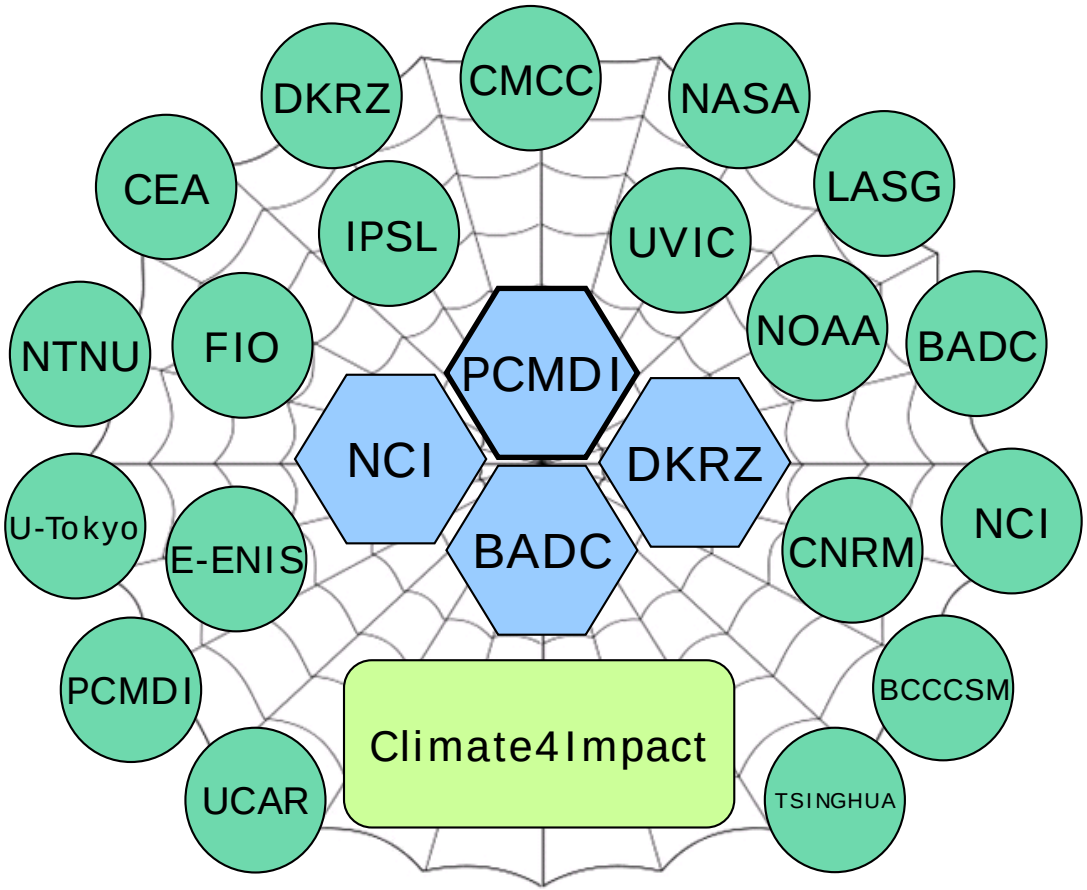


Work done on server

(de Boer & Plieger, 2014)

CC BY

Earth System Grid Federation



ESGF
Earth System Grid Federation

- Robust and distributed
- Index nodes (blue)
- Data nodes (green)
- Global Climate Model Data (CMIP5)
- Regional Climate Model Data (CORDEX)
- ~3 Petabyte of data
- Security: OpenId and X509
- **Search API offered**
- **OpenDAP data access offered**

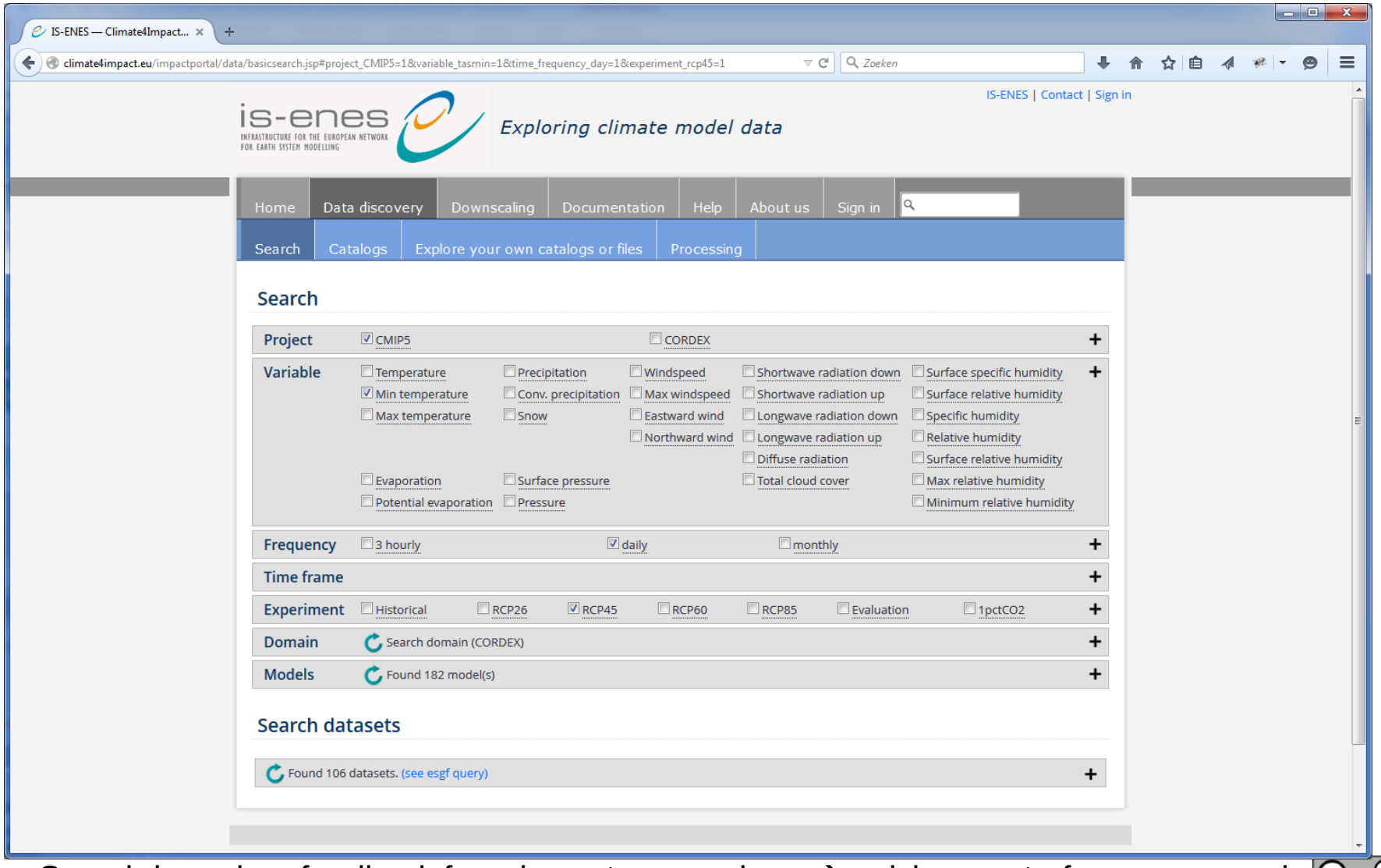
- Climate4impact builds on and contributes to this global infrastructure

OPeNDAP – for data access and subsetting

- OPeNDAP is the name of the organization and the name of the protocol
  - Open-source Project for a Network Data Access Protocol
- Data is stored at remote server
- Data model is similar to NetCDF's data model (with differences)
  - N-dimensional array container, with variables, dimensions and attributes
- Only requested pieces of data are sent
  - Accessing small pieces of large files on a remote server can still be quick
  - Data is requested based on sub-setting along dimensions
- OPeNDAP resources can be opened locally on your computer as if it were local files using the NetCDF library
  - Local files versus remote files is transparent
- The concept of a file is gone, an OPeNDAP endpoint can represent thousands of files aggregated along a dimension
  - E.g. Usually concatenate a large time series observation to one endpoint using the time dimension

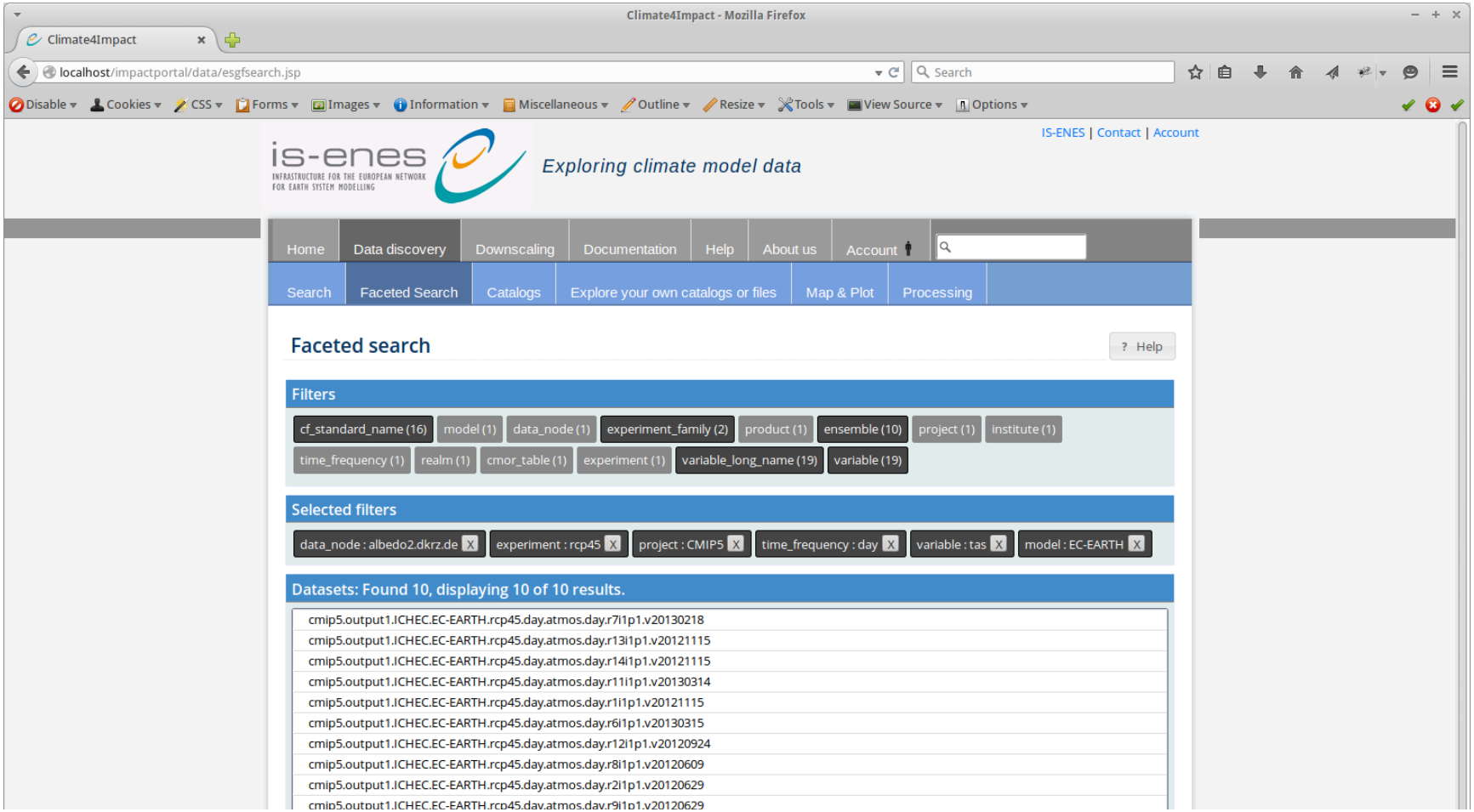
OPeNDAP within ESGF is served using the THREDDS data server

CC BY



Search based on feedback from impact researchers → quick preset of common queries

New faceted search

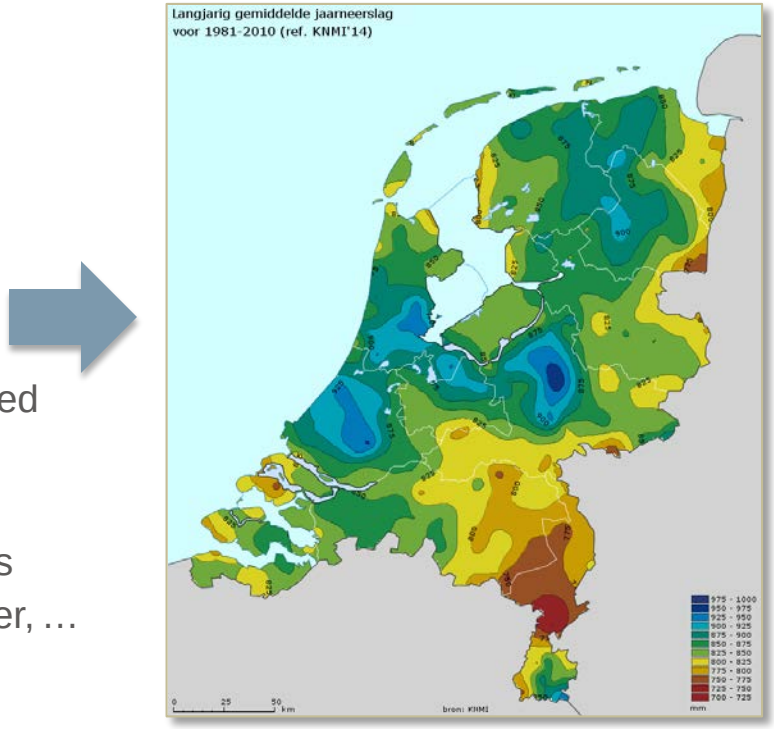


Faceted search allows to drill down search results using available filters in the federation
Results from a search query are treated as a new dataset

Web Map Service – for visualizations

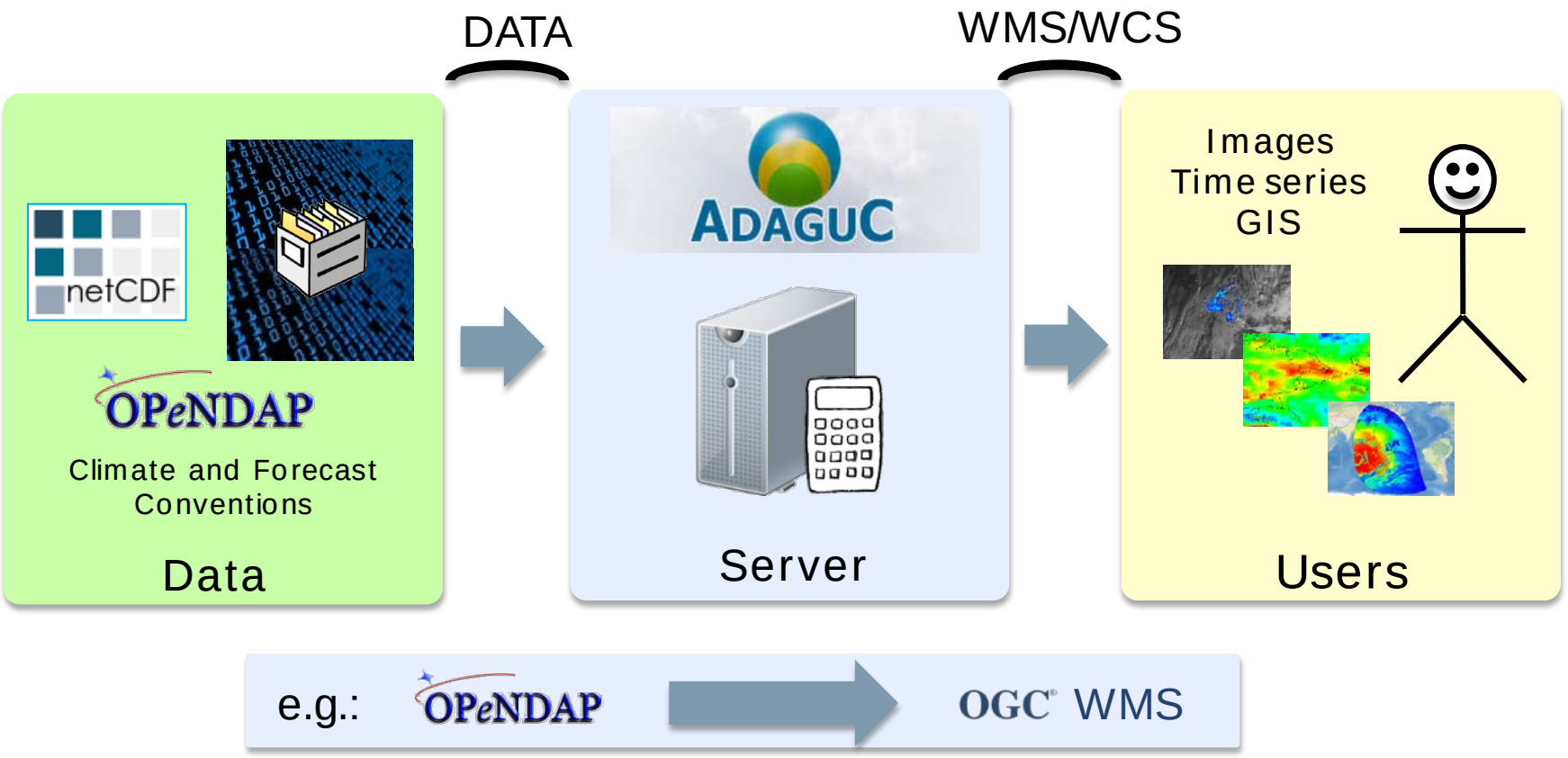
- Generates visualizations of geospatial data in the form of 2D images, suitable for transfer over the internet (JPG/PNG/GIF)

- REST based:
  - Compose an URL with key value pairs,
    and you will get an image!
- Standard is developed and maintained by the Open Geospatial Consortium
- Generated images are geo-referenced
  - Images from several sources can be easily combined
- Images have dimensions
  - Time, elevation, member
- WMS services can be viewed in many web based clients
  - OpenLayers, Leaflet, GoogleMaps, ADAGUC viewer, …



CC BY

ADAGUC Web Map and Web Coverage server

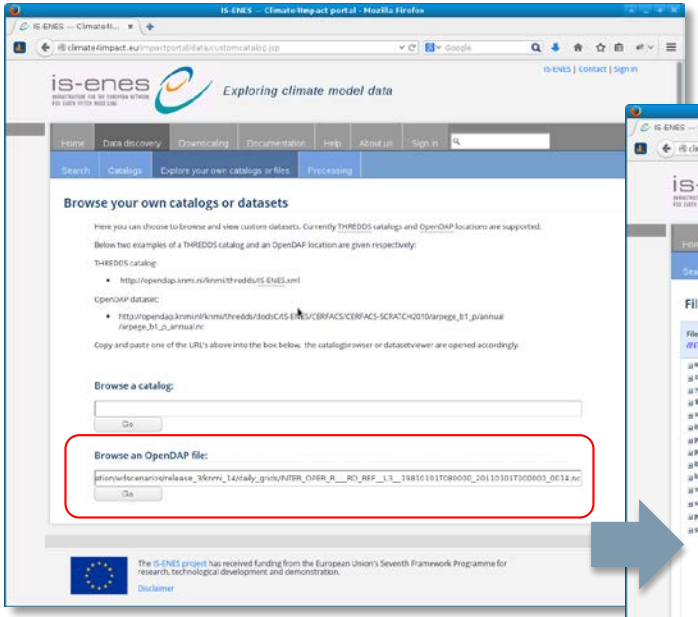


Geographical visualization framework using open standards and formats: http://adaguc.knmi.nl/

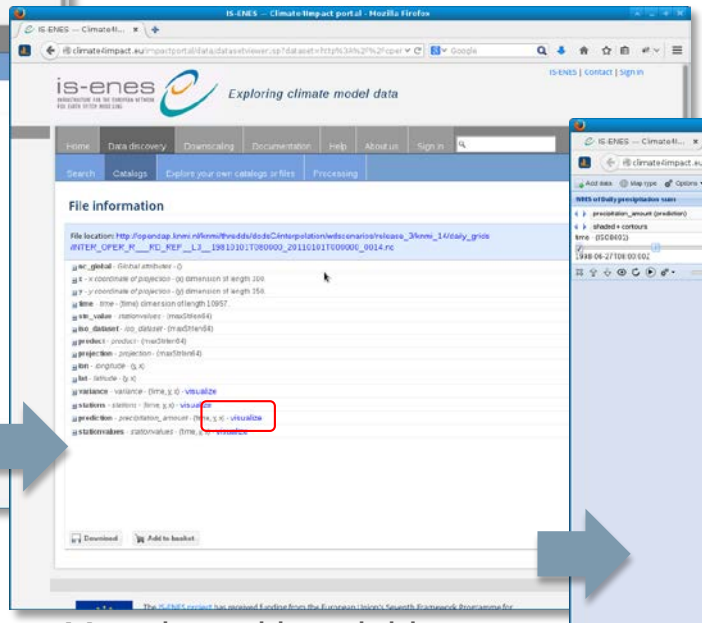
CC BY

Web Map Services based on OPeNDAP resources

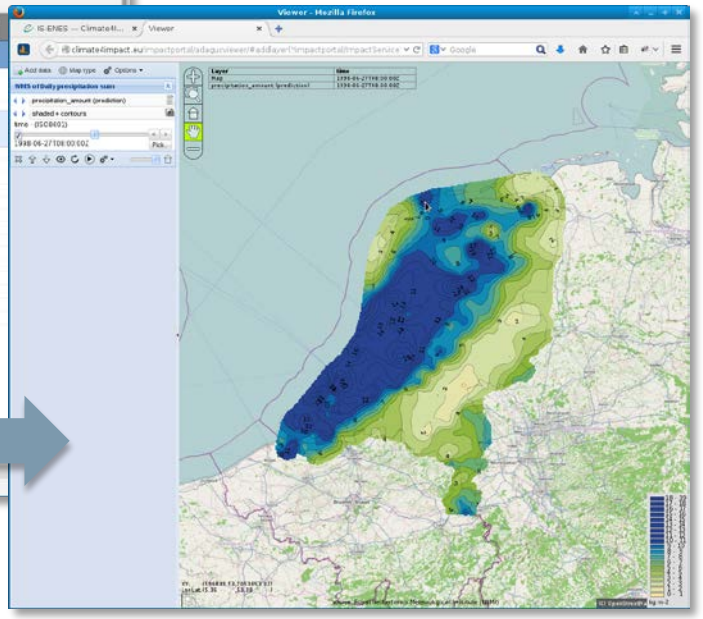
Climate4impact.eu allows for creation of WMS visualizations on OPeNDAP endpoints:



Go to “Data discovery” → “Explore your own catalogs or files”



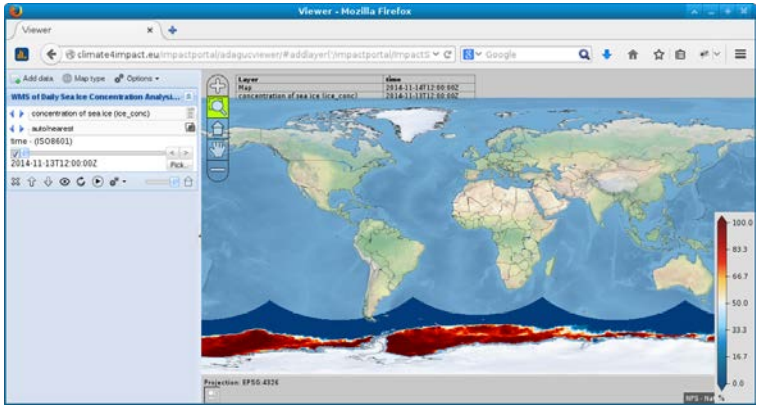
Metadata with variables to be visualized is displayed

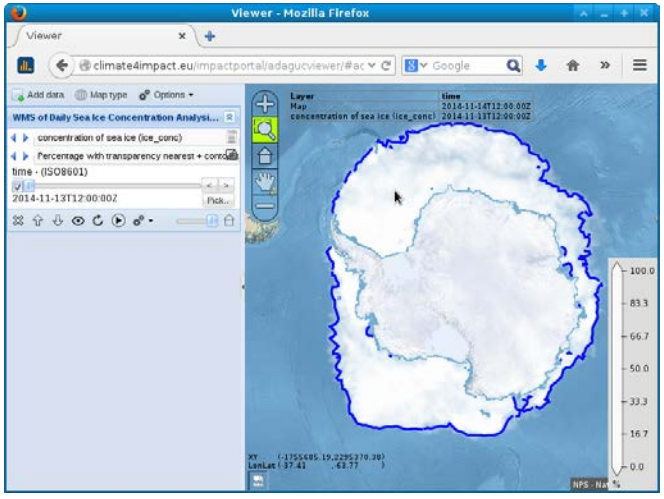


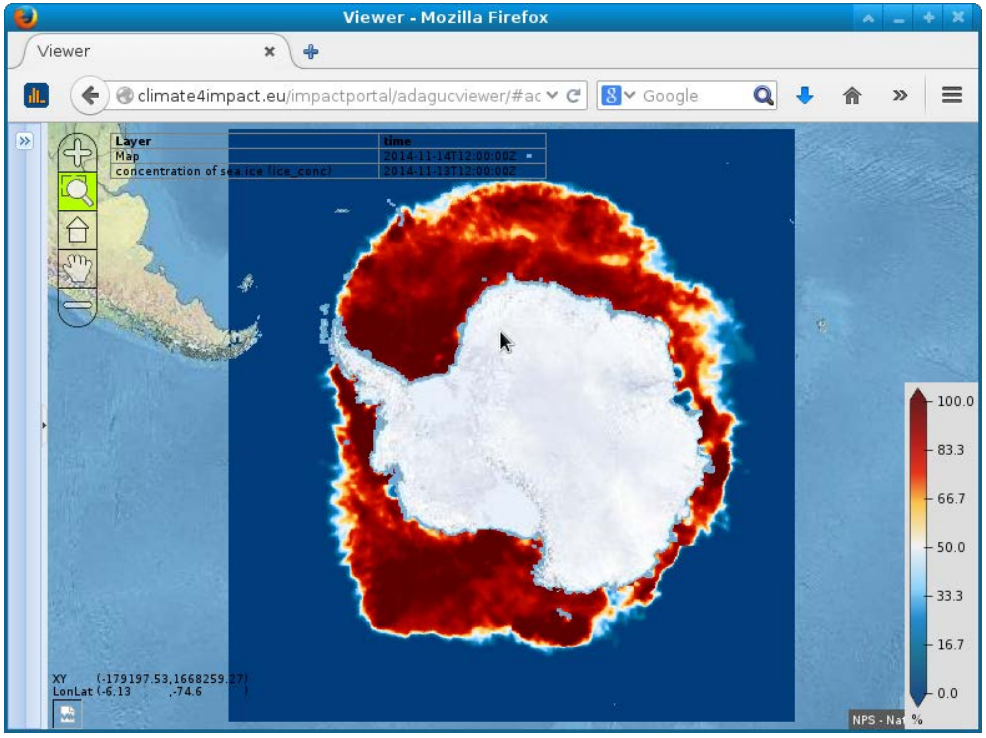
ADAGUC viewer displaying the WMS

CC BY

Example: Many existing resources with OpenDAP enabled can already be visualized!







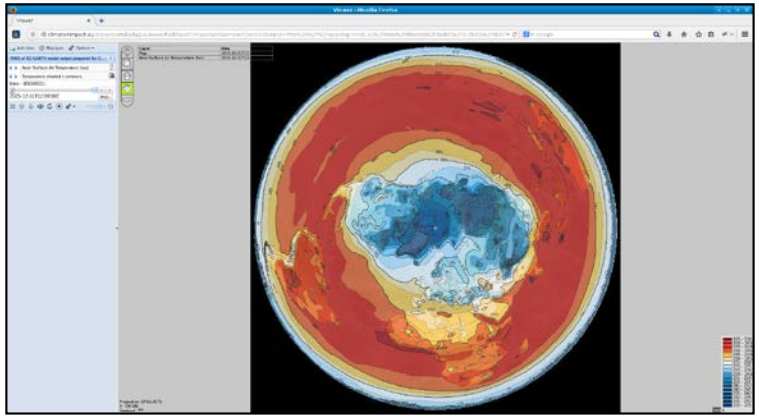
Concentration of sea ice for 2014-11-13

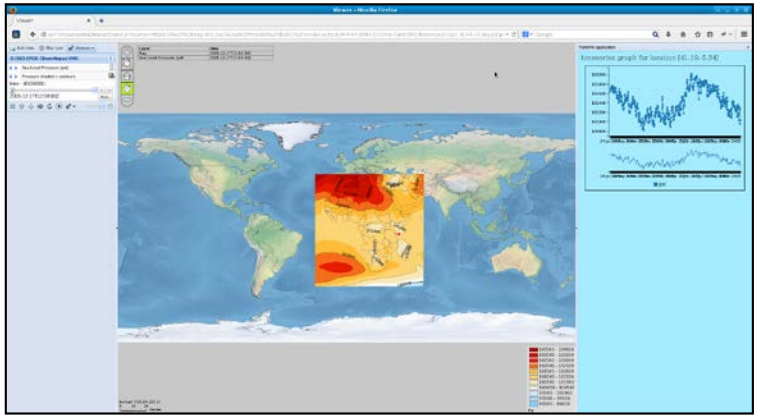
EUMETSAT Ocean and Sea ICE SAF from Norwegian Meteorological Institute

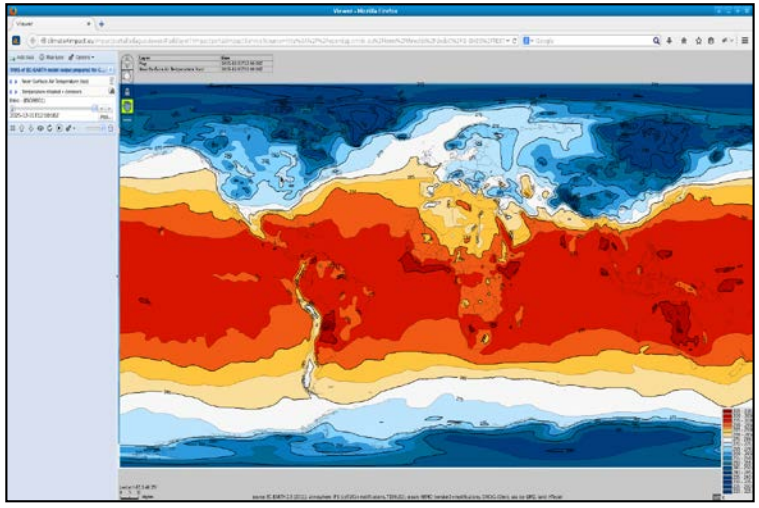
http://met.no/Hav_og_is/English/Access_to_data/

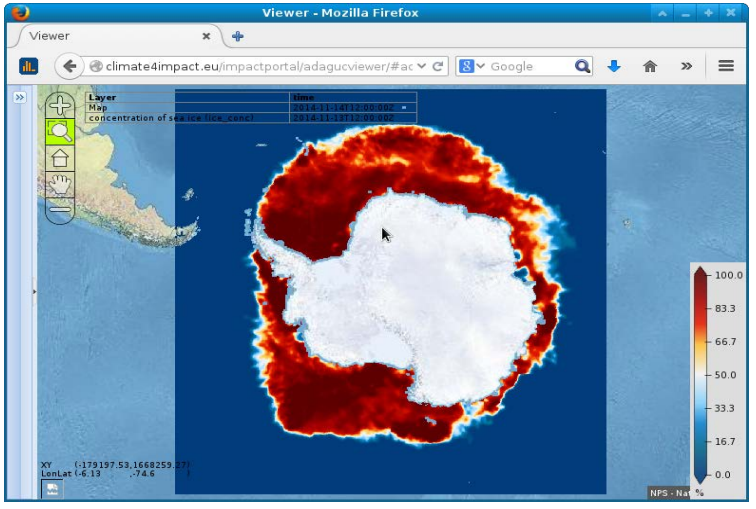
CC BY

Example: Many existing resources with OpenDAP enabled can already be visualized!









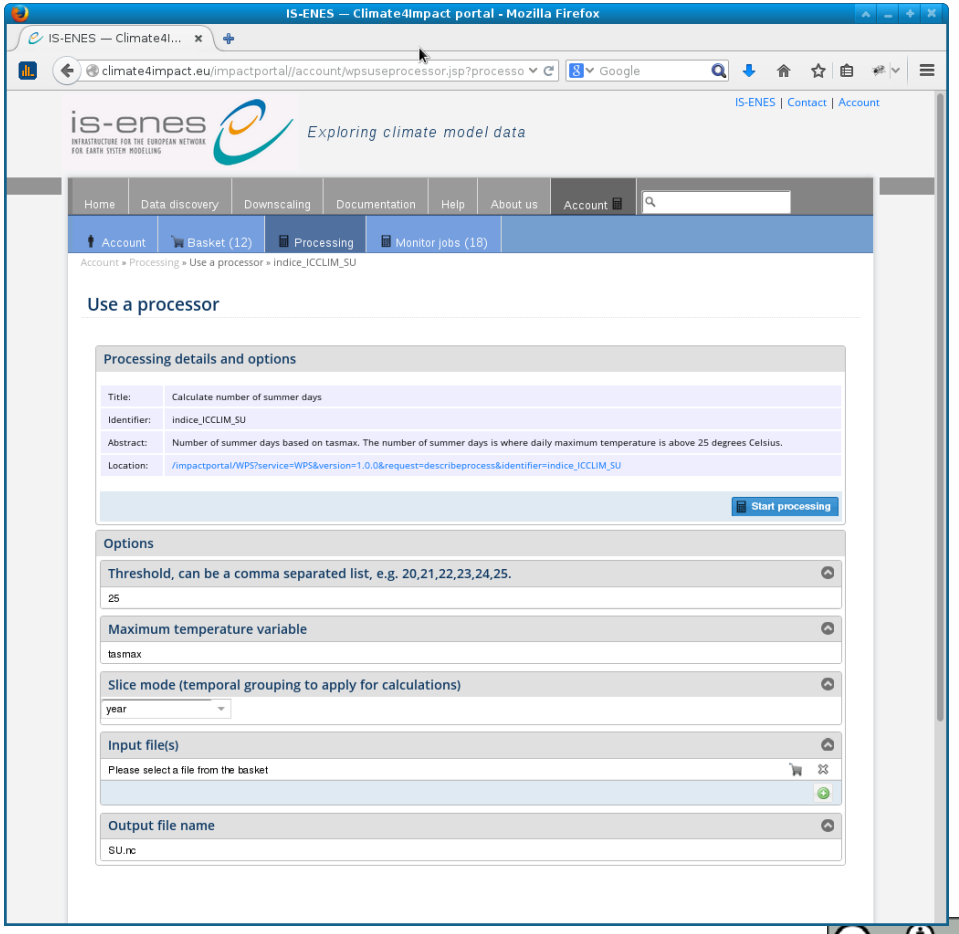
CMIP5 - global climate models

CORDEX - regional climate models

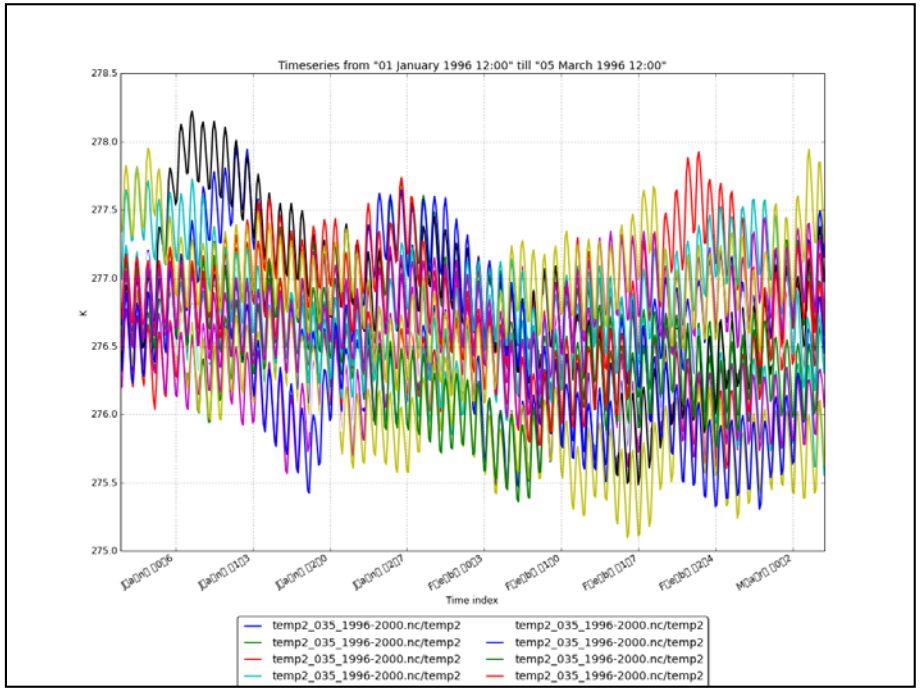
CC BY

Web Processing Service for climate indices calculations

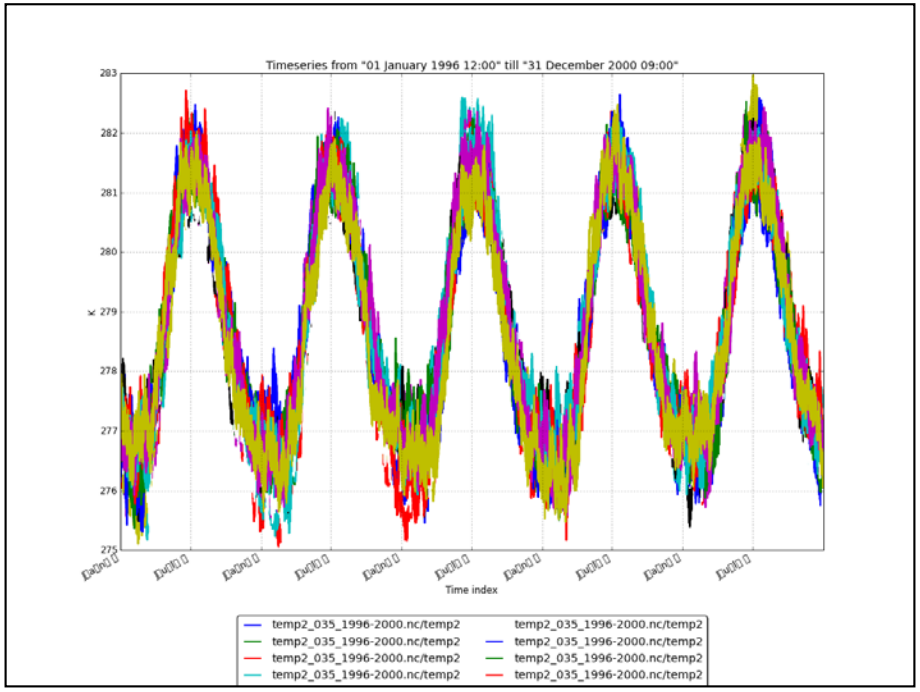
- Uses PyWPS and ICCLIM
- User interface is build automatically based on DescribeProcess XML file.
- Interface supports:
  - Link to basket
  - Comboboxes / select from list
  - Strings/text elements
- IS-ENES2 is working on an indices wizard for user friendly indices calculation
- Climate4impact WPS can be interfaced to other processing packages
- Climate indices calculations have added value!



Multi member timeseries – created with WPS



3 months

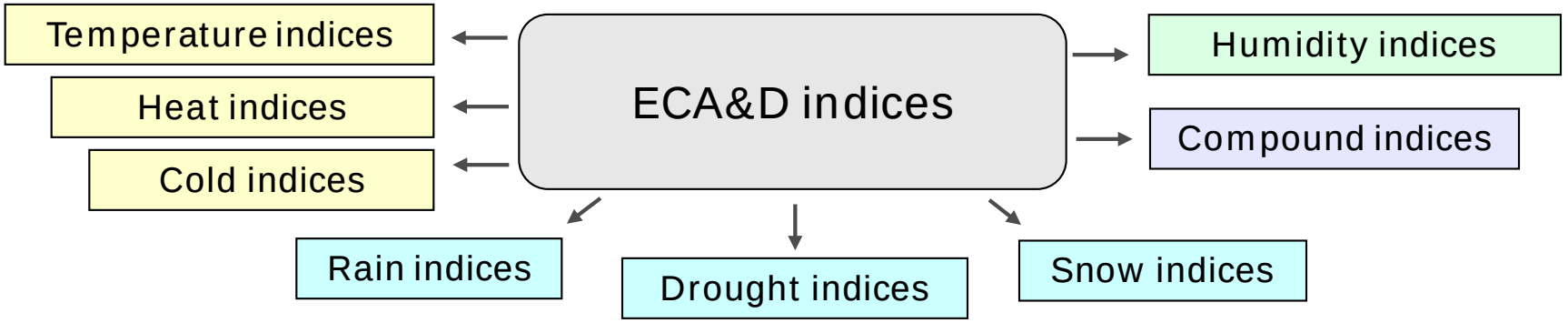


5 years

Temperature for several ensembles from the ESSENCE dataset

CC BY

Indices calculation using ICCLIM – developed in IS-ENES



- Intra-period extreme temperature range [° C] - **ETR**
- Warm days (days with mean temperature > 90th percentile of daily mean temperature) - **TG90p**
- Summer days (days with max temperature > 25 ° C) - **SU**

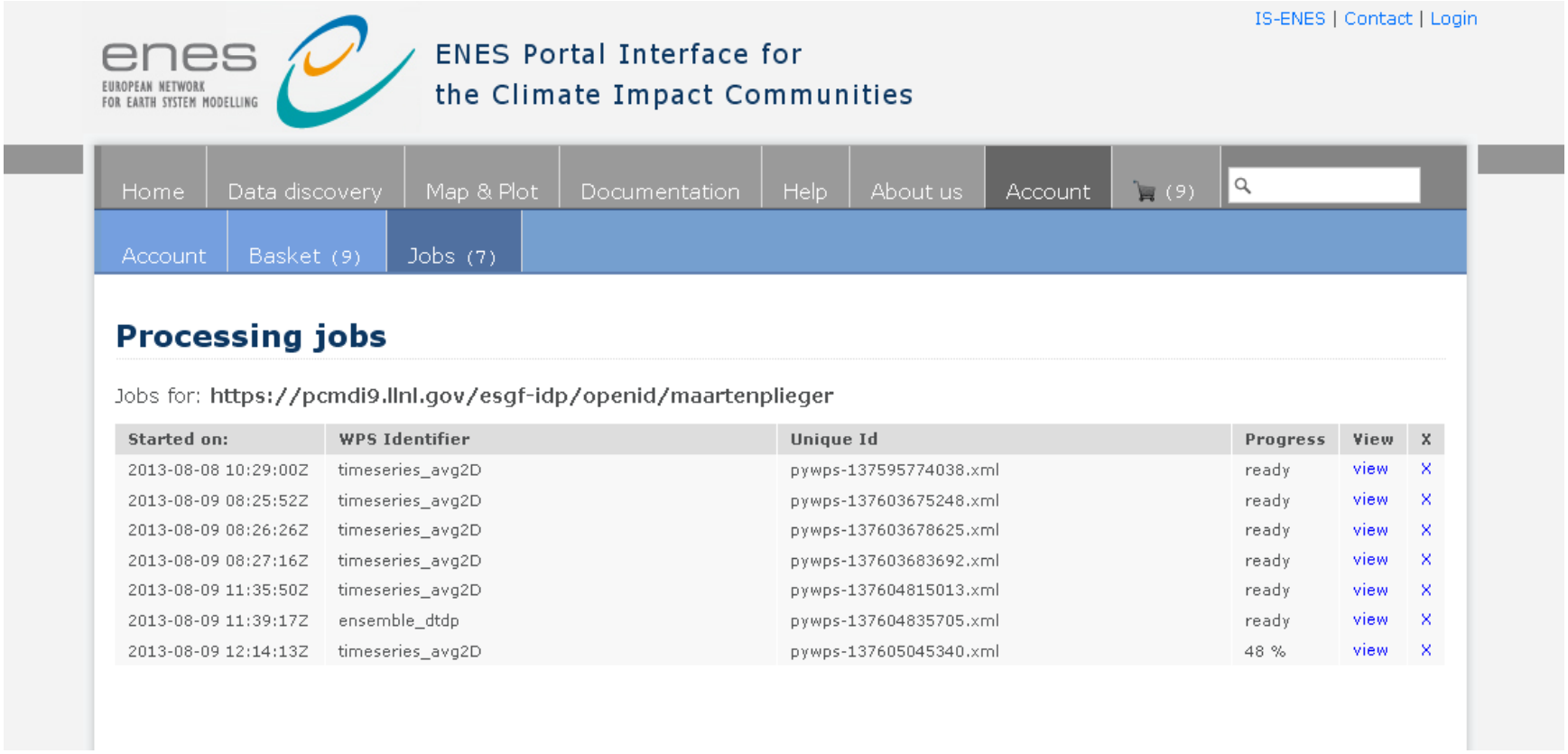
- Python code developed at CERFACS, started in September 2013
  - Generic and modular approach, can be reused in other environments
  - C functions called for optimization
- I/O interface is structured for optimal performance, with wrapper functions
- Some percentile-based indices (TG10p, TX10p, TN90p, etc): **OpenClimateGIS**

ICCLIM source code and documentation is available via https://github.com/cerfacs-globc/icclim

CC BY

Web Processing Service

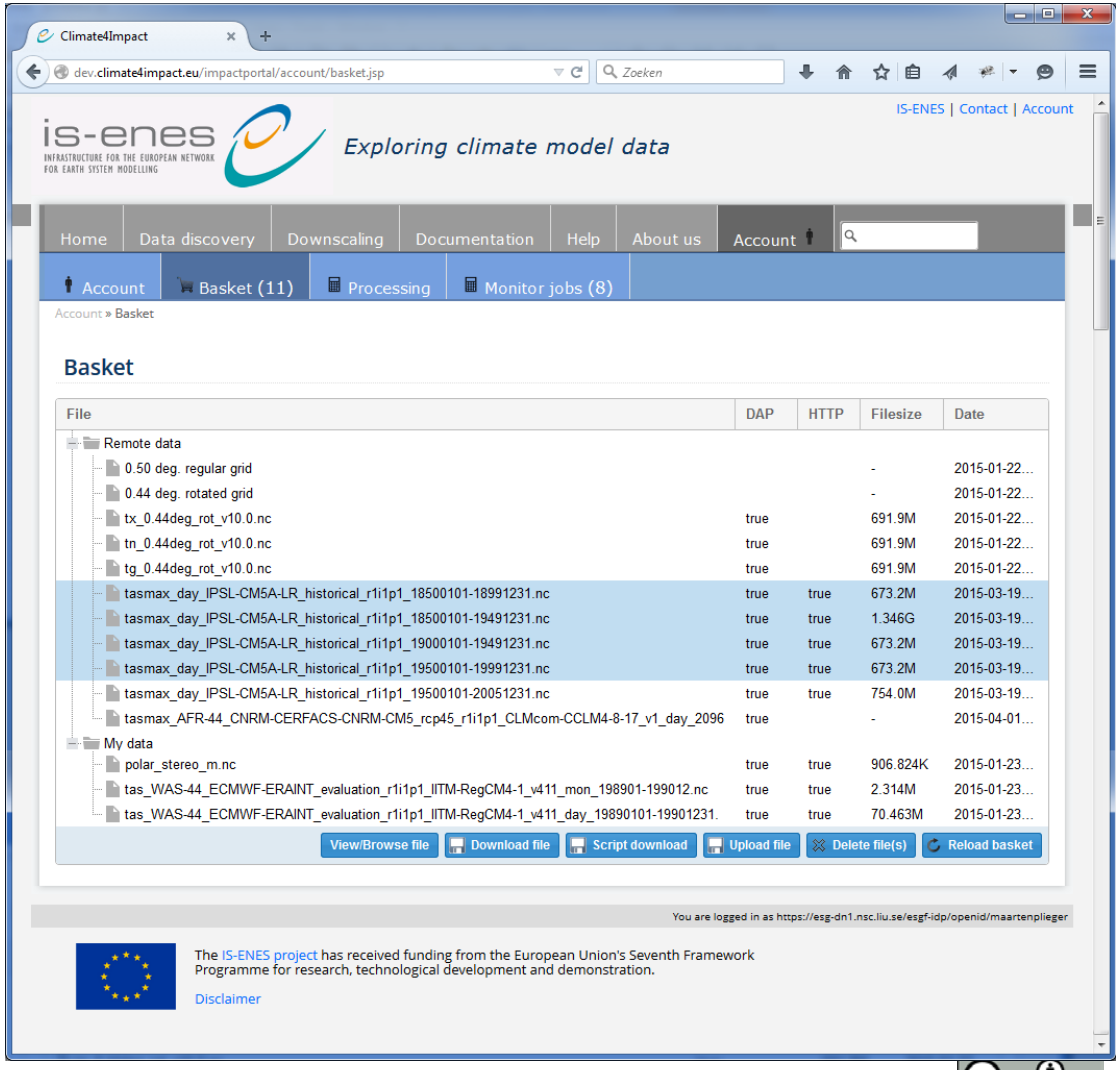
Job progress can be viewed from anywhere



Started on:	WPS Identifier	Unique Id	Progress	View	X
2013-08-08 10:29:00Z	timeseries_avg2D	pywps-137595774038.xml	ready	view	X
2013-08-09 08:25:52Z	timeseries_avg2D	pywps-137603675248.xml	ready	view	X
2013-08-09 08:26:26Z	timeseries_avg2D	pywps-137603678625.xml	ready	view	X
2013-08-09 08:27:16Z	timeseries_avg2D	pywps-137603683692.xml	ready	view	X
2013-08-09 11:35:50Z	timeseries_avg2D	pywps-137604815013.xml	ready	view	X
2013-08-09 11:39:17Z	ensemble_dtdp	pywps-137604835705.xml	ready	view	X
2013-08-09 12:14:13Z	timeseries_avg2D	pywps-137605045340.xml	48 %	view	X

CC BY

- By default the basket contains:
  - "Remote data" for links
  - "My data" for your own data
- Script based download allows to select and download multiple files
- Client certificate (x509) is embedded in download script
  - No need for MyProxy login
  - No need for firewall changes
- The basket allows for uploading your own files
  - Can be used in processing or visualization



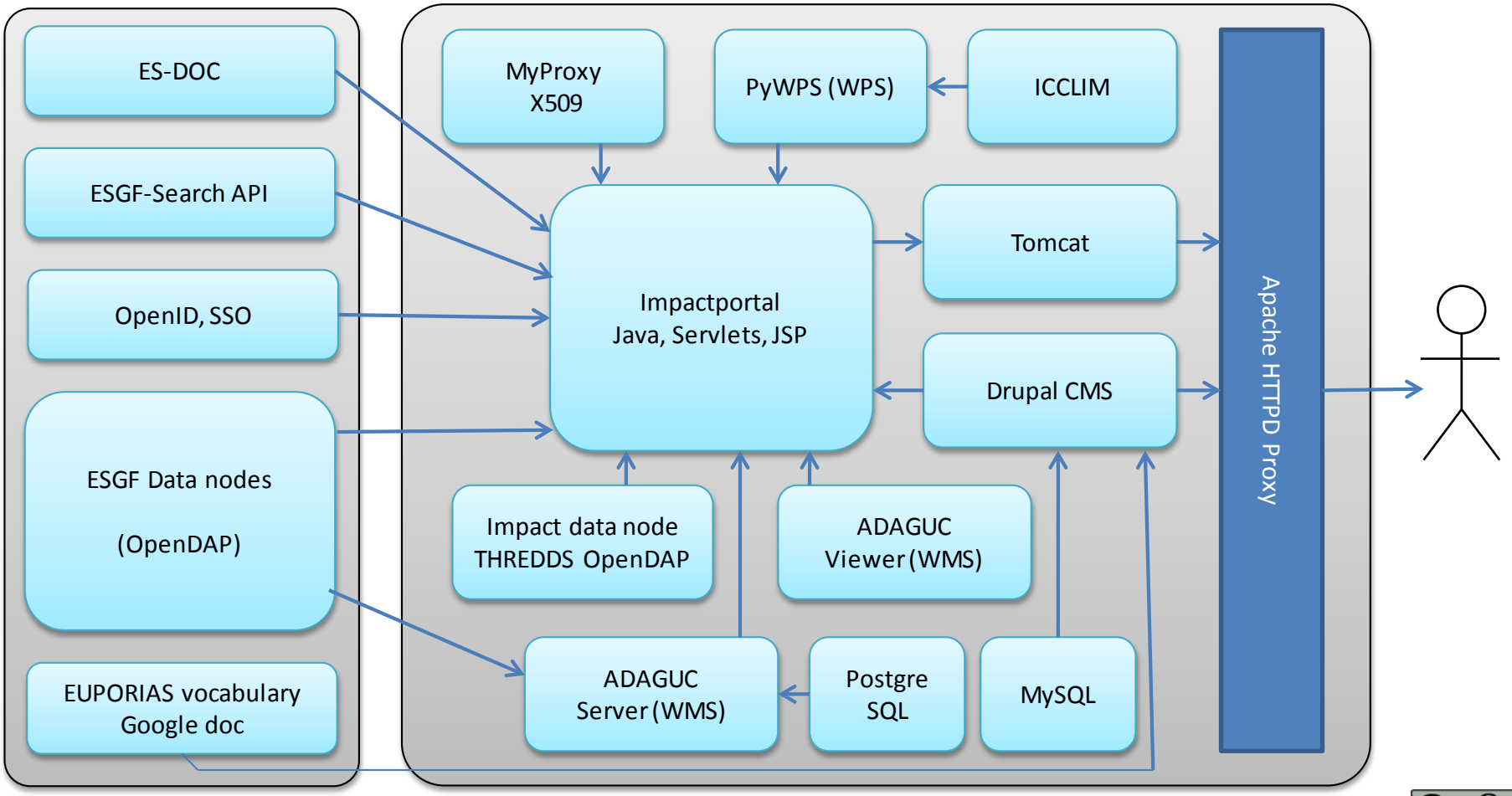
Conclusion and next steps

- Climate4impact enables impact research

- Climate4impact is flexible due to applied technologies and standards
  - ADAGUC WMS can be used to visualize local and remote files
  - PyWPS with ICCLIM is suitable as generic processing framework for climate indices
  - OPeNDAP can be used to access small bits of large files over the internet quickly

- Climate4Impact enables easier use of ESGF services
  - Search many datasets from several projects
  - Data access and visualization via OPeNDAP
  - Security is offered in a more user friendly way

- Next steps:
  - Improve user interface → make more user friendly
  - Implement use cases from climate impact researchers
  - Finalize the connection to the University of Cantabria downscaling portal
  - Climate indices calculation wizard

CC BY

How components fit together