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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°824084

climate4impact.eu (2.0)

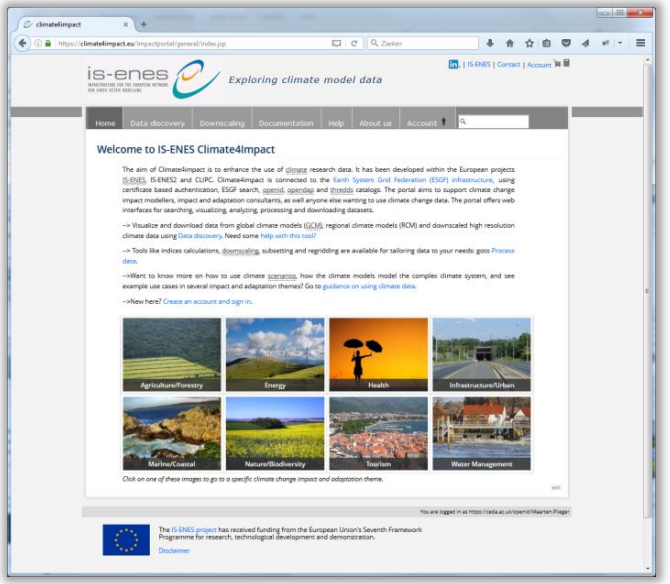
Enhance usage of research data and support researchers with climate analysis. Current state and Next steps

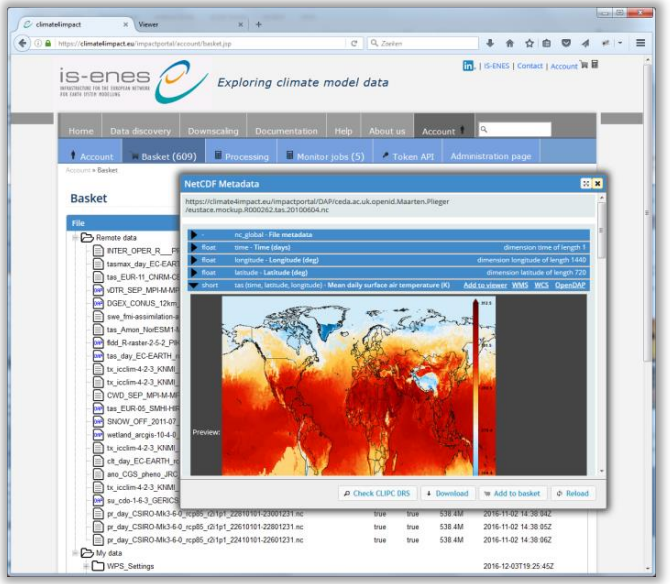
Christian Pagé, **Maarten Plieger**, Wim Som de Cerff, Alessandro Spinuso, Ernst de Vreede, Niels Drost

KNMI, CERFACS, Netherlands eScience Center

What is the climate4impact portal?

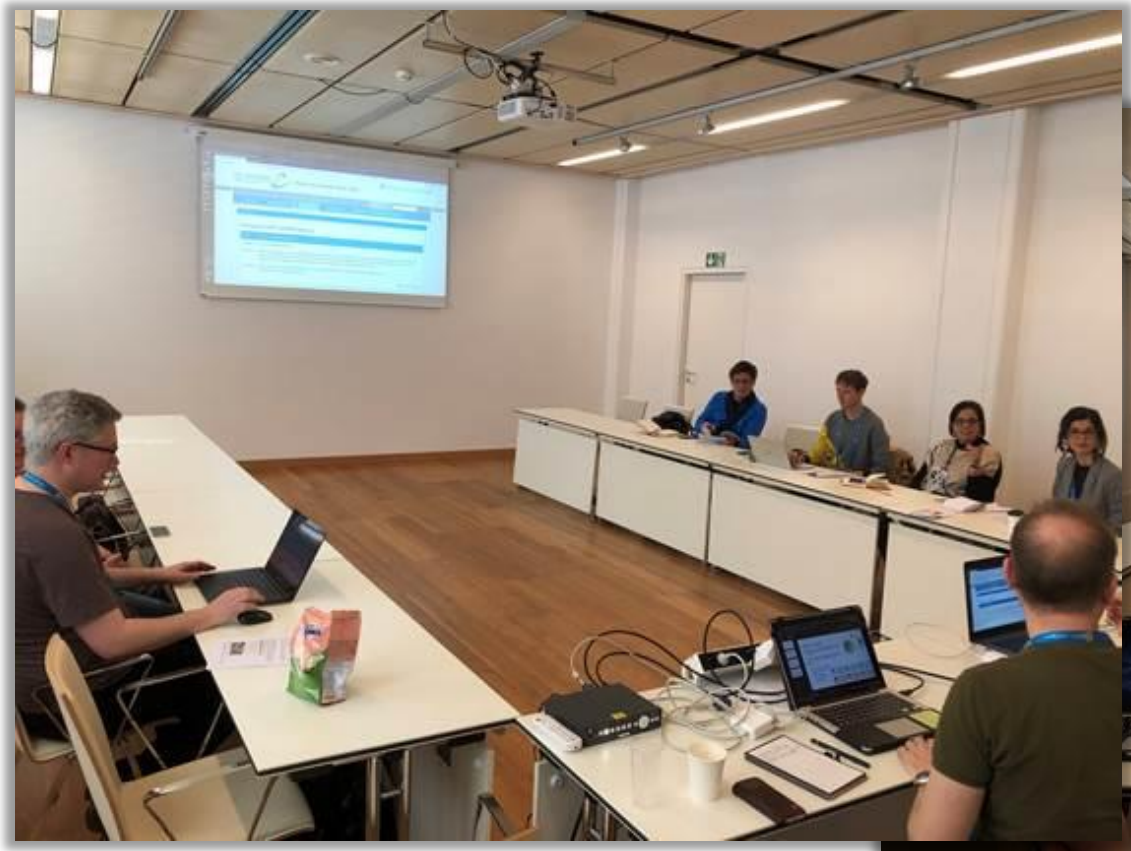
- Platform for researchers to explore climate data and perform analysis
- Connects to ESGF web services
  - ESGF search, THREDDS support, Security
  - CMIP6, CMIP5, CORDEX, other MIPS
- Visualization via ADAGUC
  - Visualization system using Web Map Services
  - Web Coverage Services for data transformation
- Analysis using (Py)WPS to perform calculations
  - ICCLIM climate indices calculation, data reduction
  - Personal store for processing outcomes
- In production
  - Deployed in the cloud and is used
  - Will be one of the CMIP6 dissemination portals





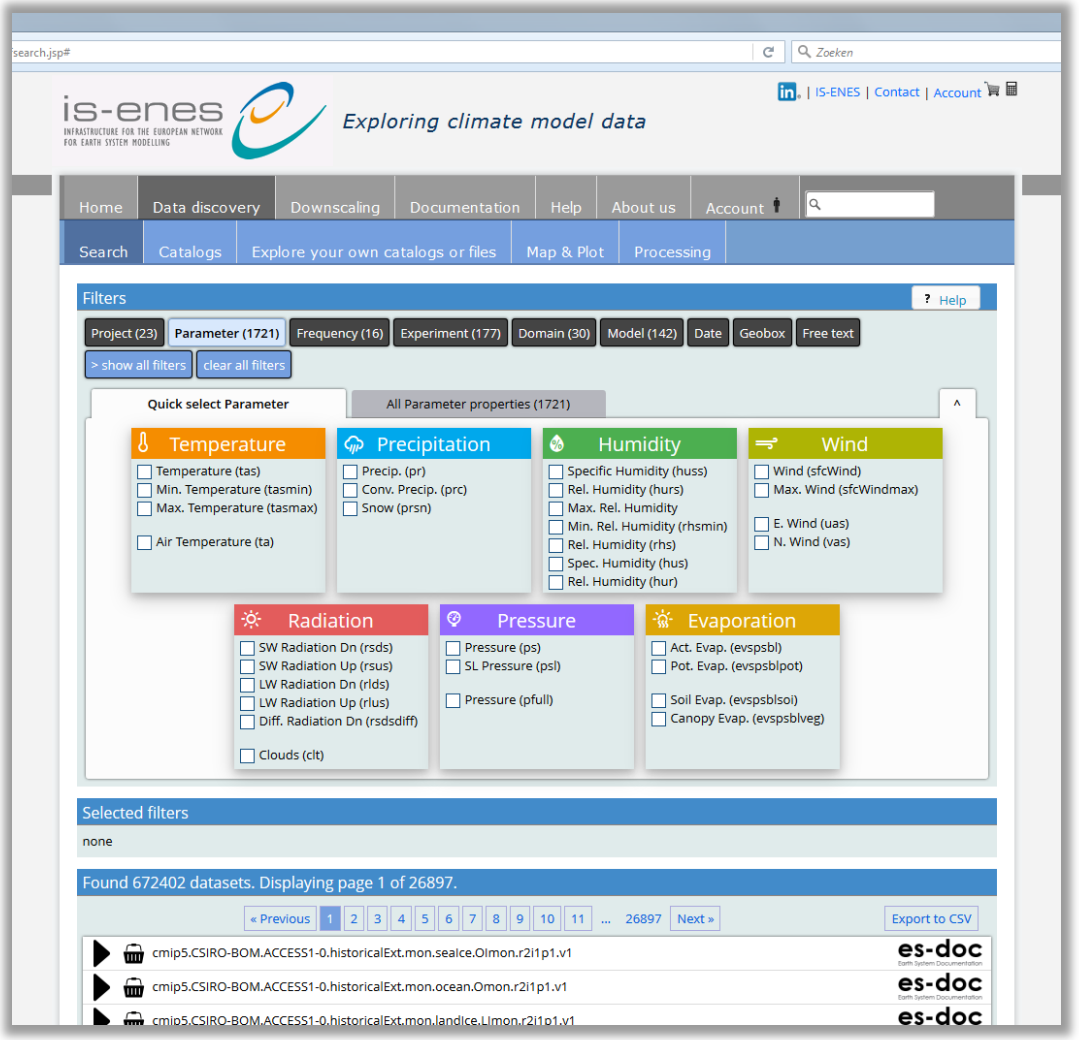
Splinter sessions, web seminars, classrooms

- Receive feedback from users
- Keep users in the loop, show them new features



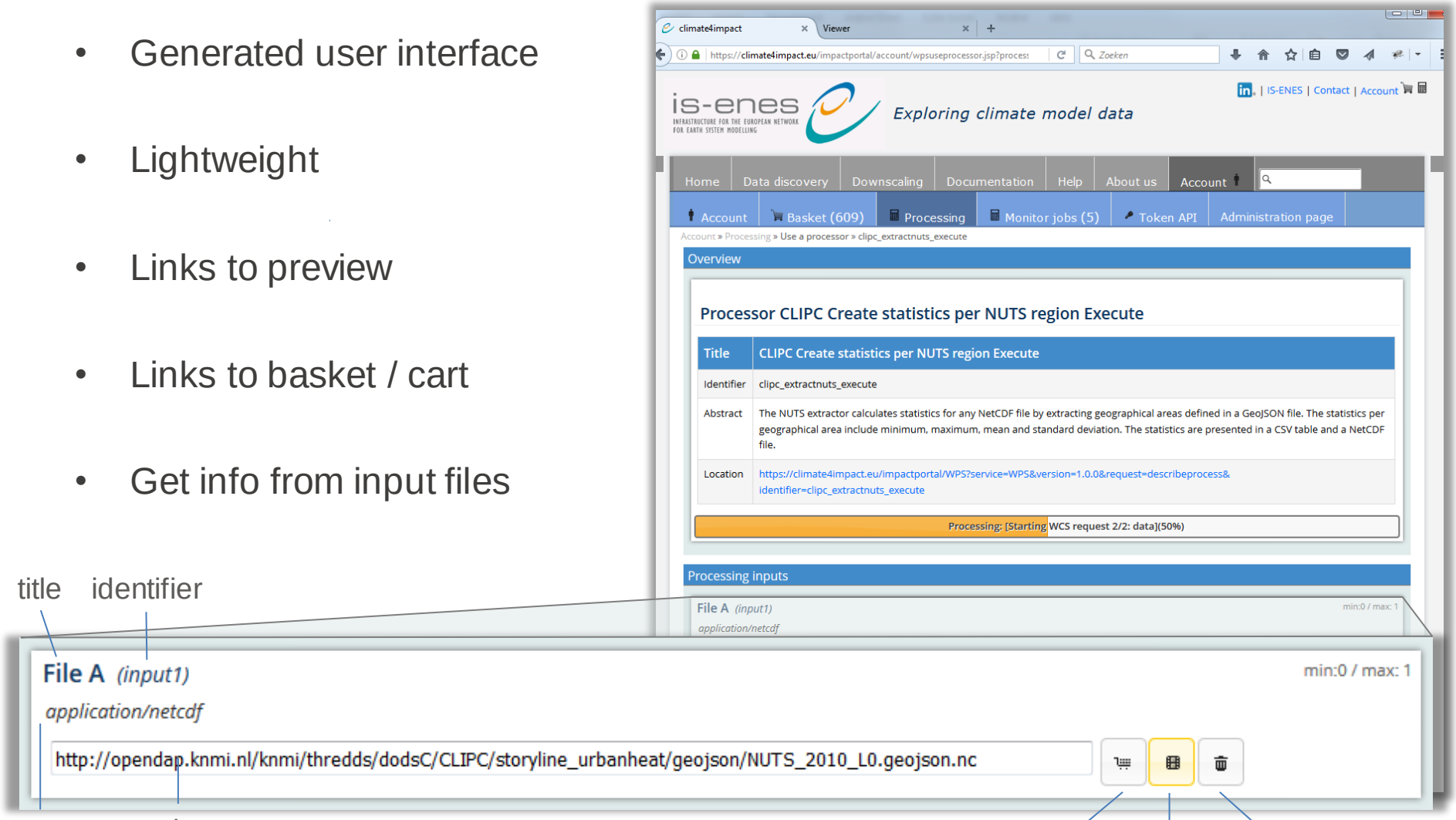
Web based faceted search for any (climate) data via ESGF

- Drill down search results
- Tooltips for acronyms
- Quick select menus
- ES-DOC integration
- Preview of data
- Export search list to CSV



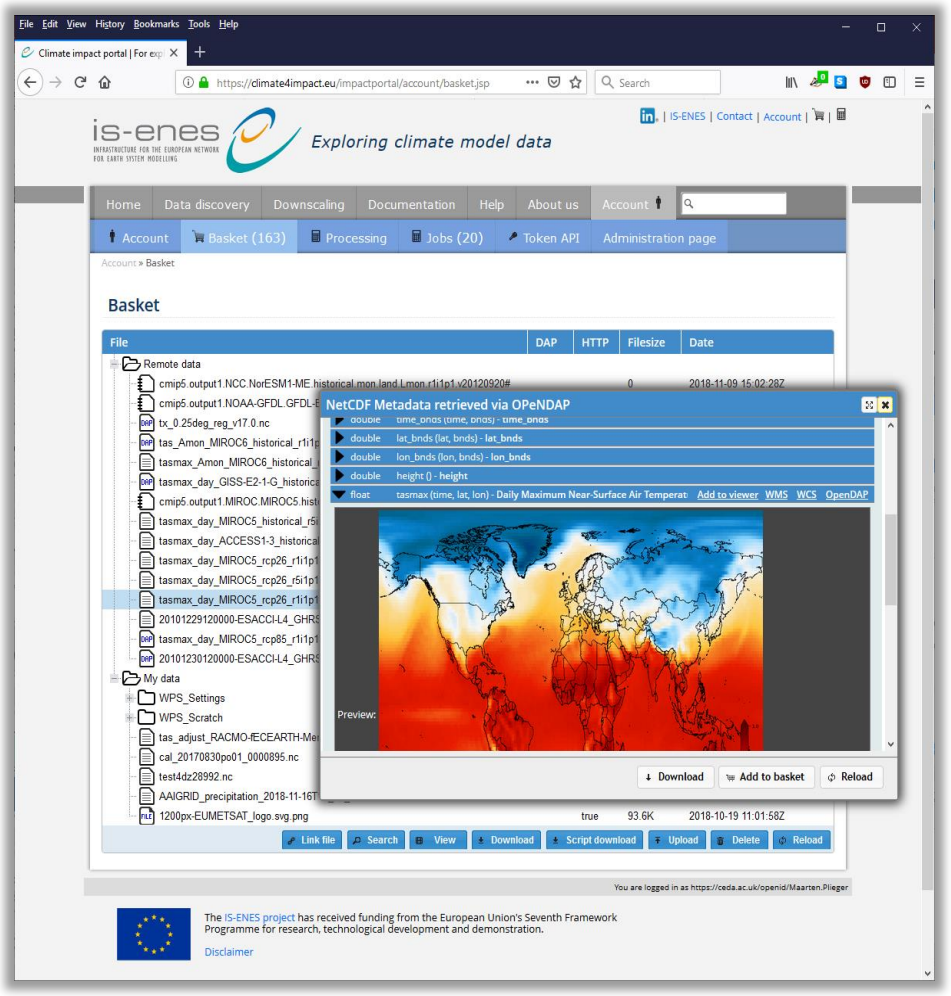
Web processing interface for your own data analysis

- Generated user interface
- Lightweight
- Links to preview
- Links to basket / cart
- Get info from input files



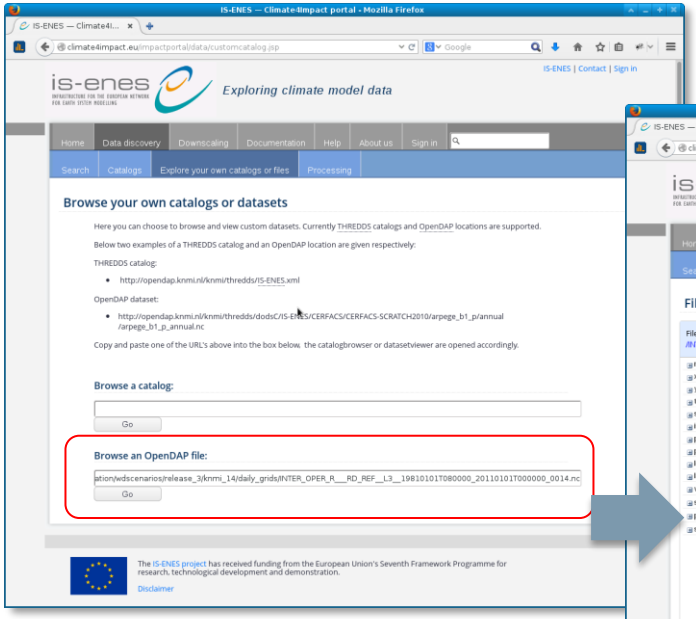
Personal basket

- 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
- The basket allows for uploading your own files
  - Can be used in processing or visualization
  - NetCDF, CSV, GeoJSON, PNG
- Share your data located in your basket with others

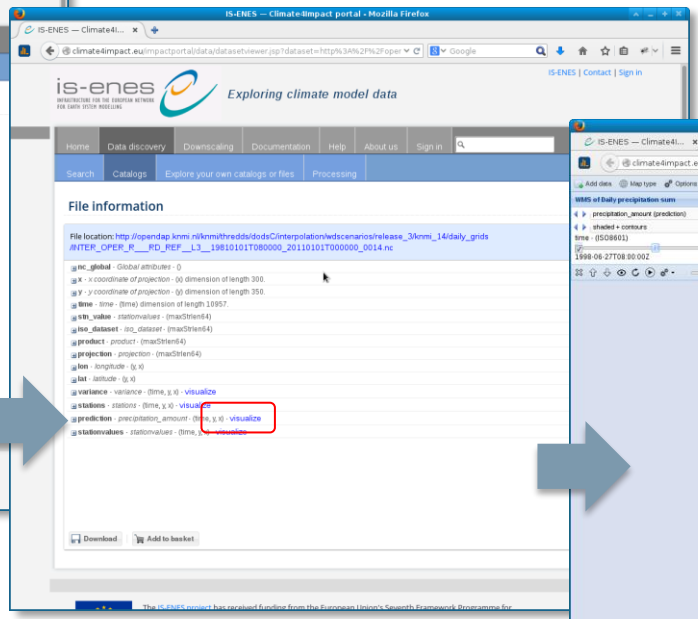


Visualization of (any) THREDDS catalogs and OPeNDAP

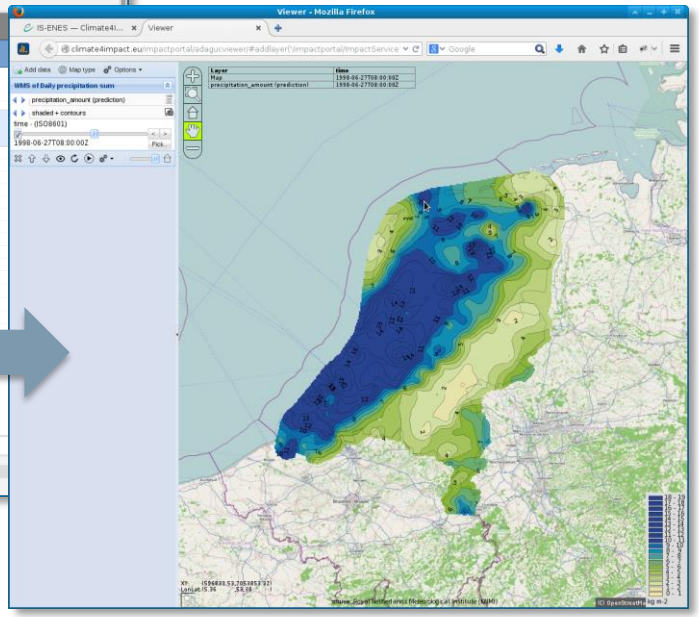
You can connect to remote datasets, not only data from the ESGF



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

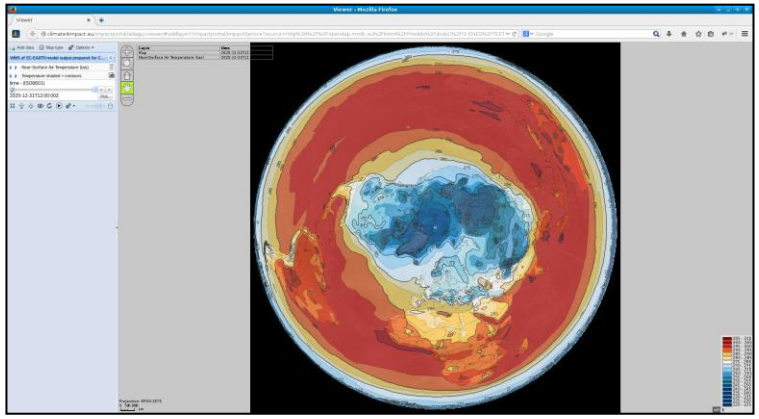


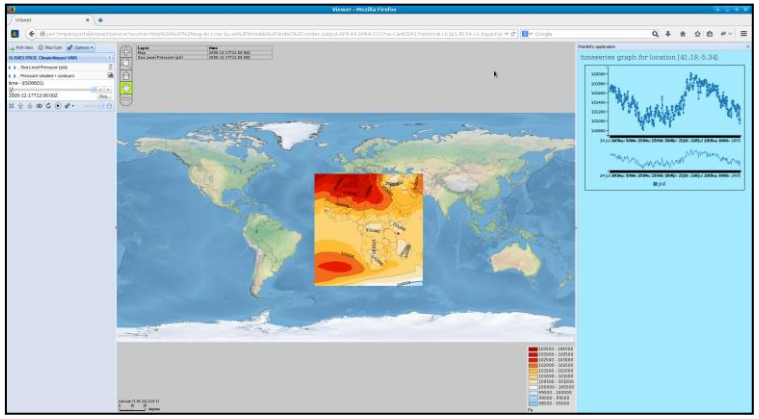
Metadata with variables to be visualized is displayed

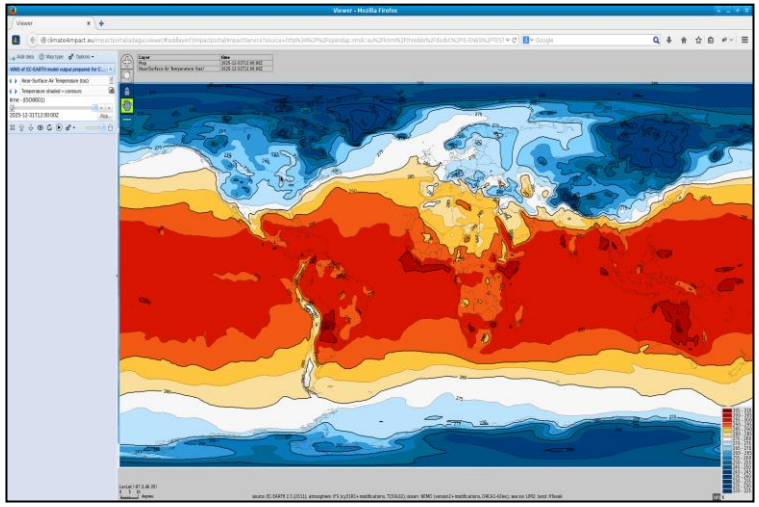


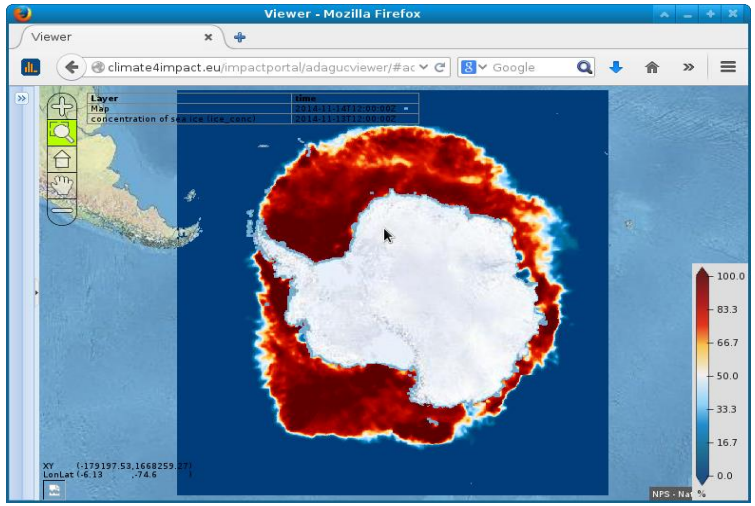
ADAGUC viewer displaying the WMS

Visualization – using ADAGUC web mapping tools









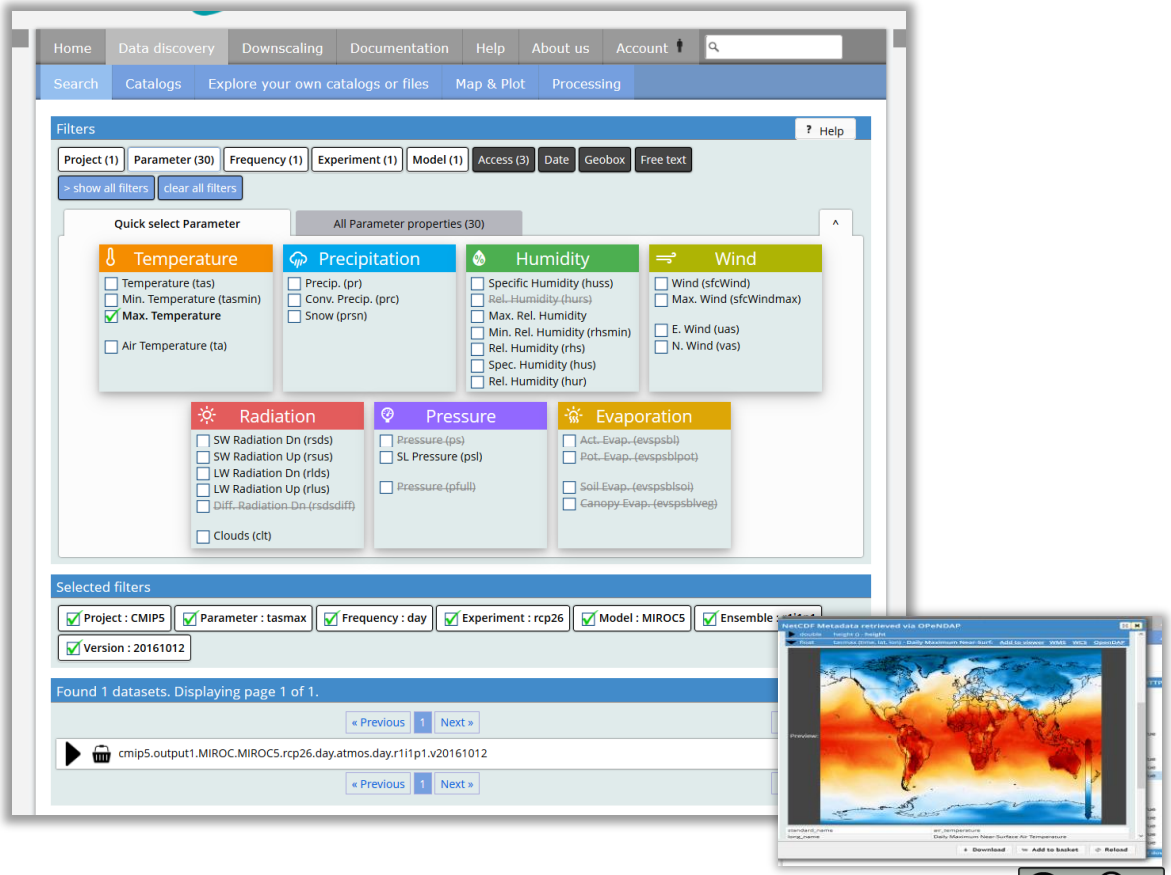
CMIP5 - global climate models

CORDEX - regional climate models

Example use case: Calculating summer days (SU) - step 1

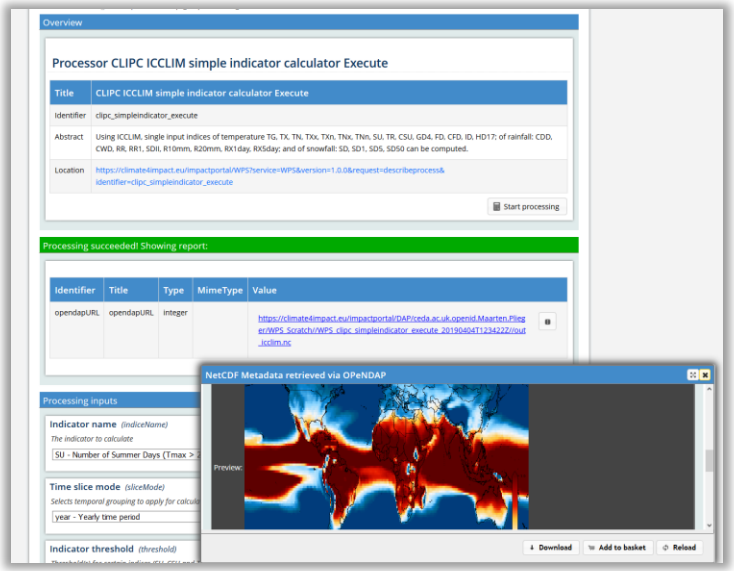
- Calculate number of days where maximum temperature is above 25 degrees per European country, based on experiment RCP 2.6 and model MIROC5

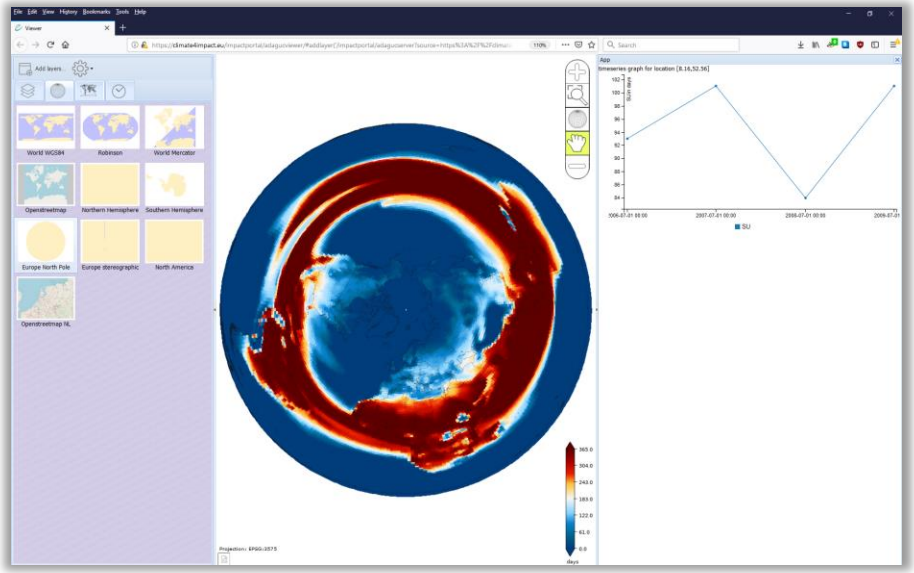
- Go to C4I and sign in
- Go to Search and select:
  1. Model: CMIP5
  2. Parameter: tasmax
  3. Time frequency: daily
  4. Experiment: rcp26
  5. Model: MIROC5,
  6. Ensemble: r1i1p1
  7. Select the latest version
- Select a file from the dataset and add it to your basket



Calculate Summer days – step 2

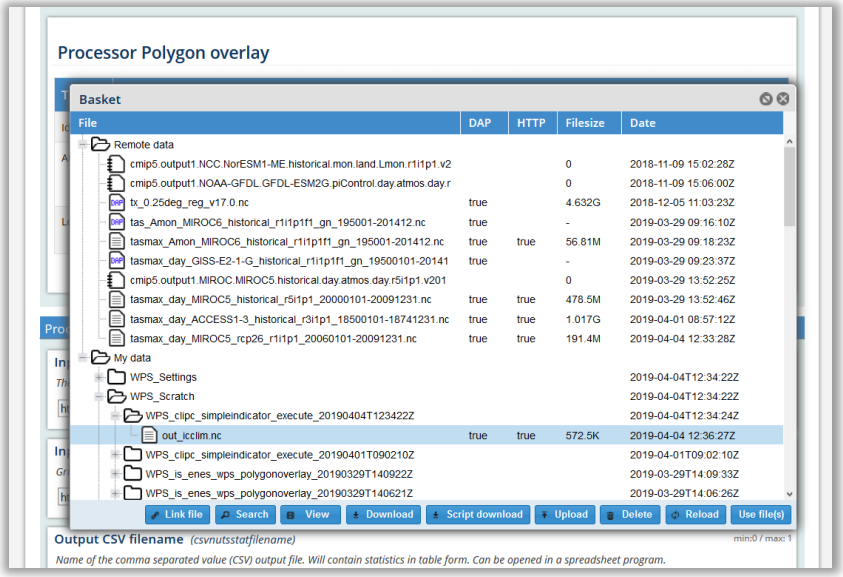
- Go to Processing and select ICCLIM simple indicator calculations
- Select SU, Summer days. Leave the threshold to 25 degrees Celsius
- Select the file from your basket and click “Start processing”
- Inspect the output

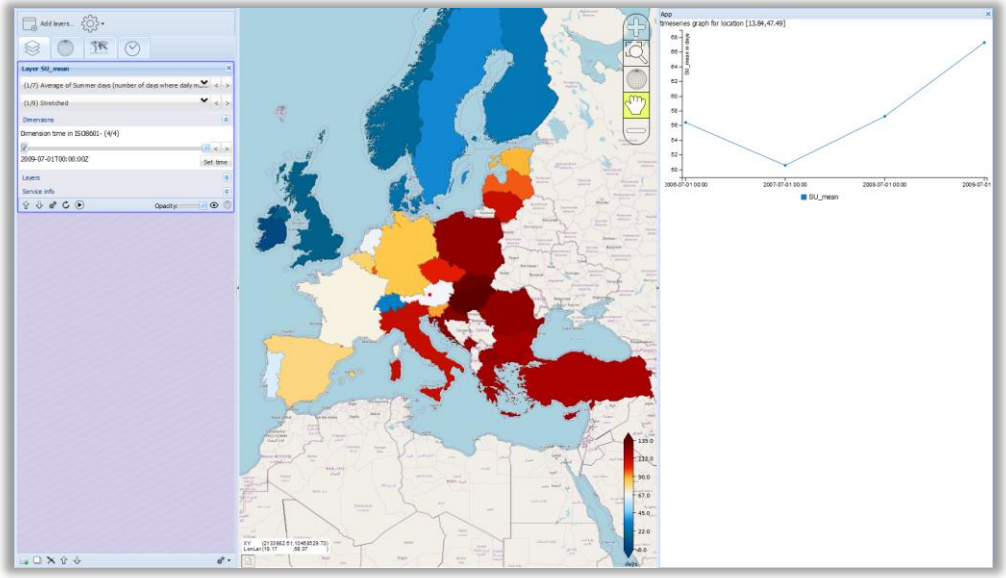




Calculate region statistics – step 3

- Go to Processing and select Polygon overlay
- For “Input File B - Gridded data”, choose the latest result with SU from your basket. This is the most recent folder under WPS_Scratch
- As variable select “SU”, as time range select “*”
- Click “Start processing”





Result: Summer days per European country for MIROC5 / RCP26 !

What can be improved? – Next steps in IS-ENES3

- Currently C4I handles ESGF data on file level
  - X Fragmentation of files is a barrier for many users and hurts user experience
  - ➔ Hide file structure, work with datasets and search patterns
  - ➔ Especially important for **CMIP6,** C4I will become one of the data distribution endpoints for CMIP6

- Currently the Processing services are on the same machine as the portal
  - X Currently not scalable and processing load effects the portal
  - ➔ Make use of distributed Web Processing Services
  - ➔ E.g. DKRZ's BirdHouse WPS framework, ESGF CWT and results from H2020 DARE
  - ➔ Adaguc-services framework allows execution of remote WPS services

- Currently the frontend uses old technologies (JSP, Jquery, ExtJS)
  - X Difficult to maintain, and it is difficult to re-use results from other work
  - ➔ Migrate to ReactJS (Based on work done in C3S-Magic)
  - ➔ Good moment to re-design the user interface in collaboration with users.

- Currently the viewer is running in a separate tab
  - – By using ReactJS, it is easier to make use of an embedded viewer (adaguc-webmapjs)

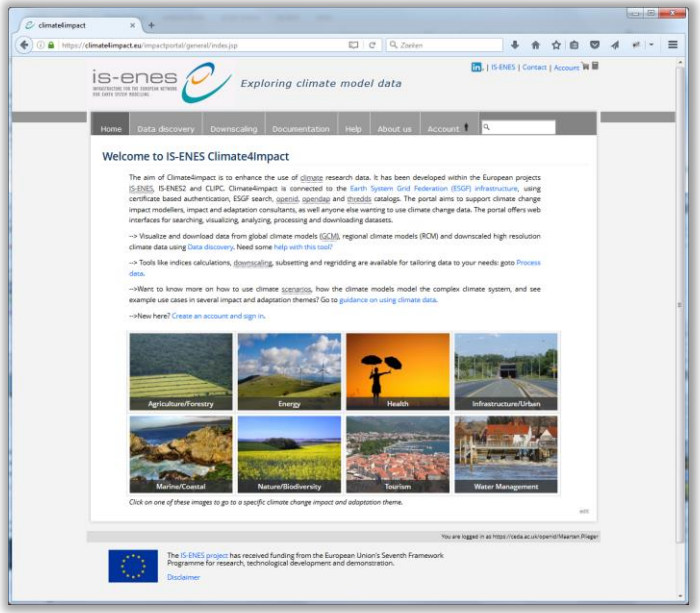
- Currently provenance tracking is limited to a few processes
  - – Enhance usage of W3C PROV-DM standard and WPS_PROV toolkit

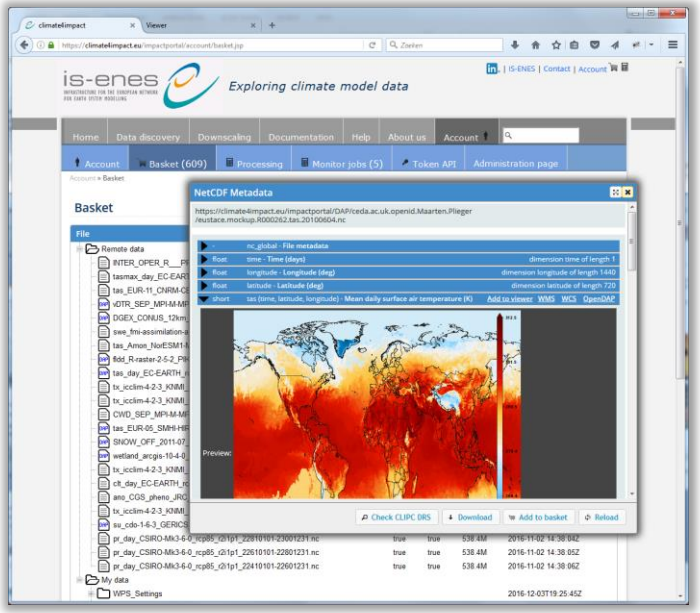
We are looking for users who are willing to help to improve the platform!

Thank you for your attention!

The climate4impact team

- Check out our ADAGUC poster on Friday 8:30-10:15, location X1.62
- For questions, suggestions, feedback and help, please contact
  - maarten.plieger@knmi.nl
  - wim.som.de.cerff@knmi.nl
  - christian.page@cerfacs.fr





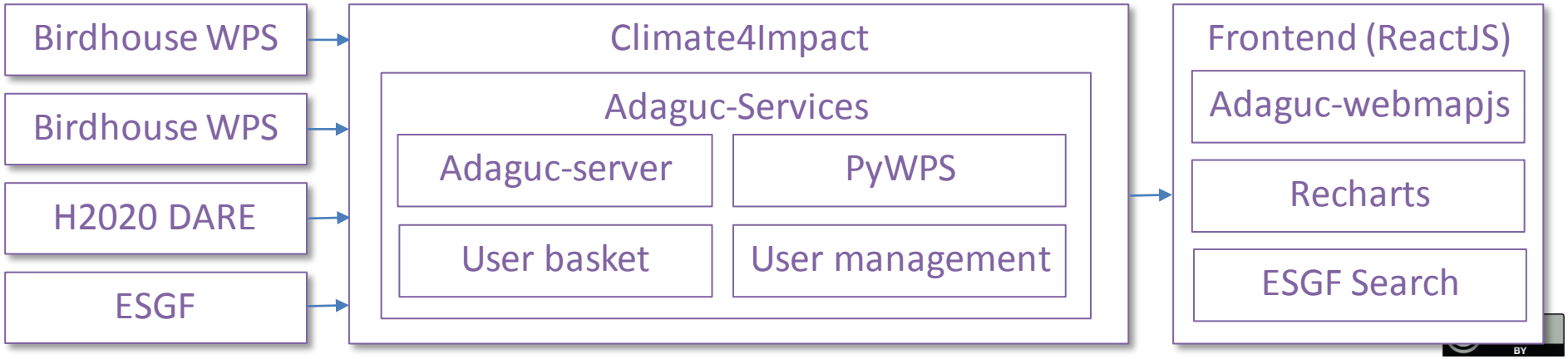

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK
FOR EARTH SYSTEM MODELLING


EGU
Photography
encouraged

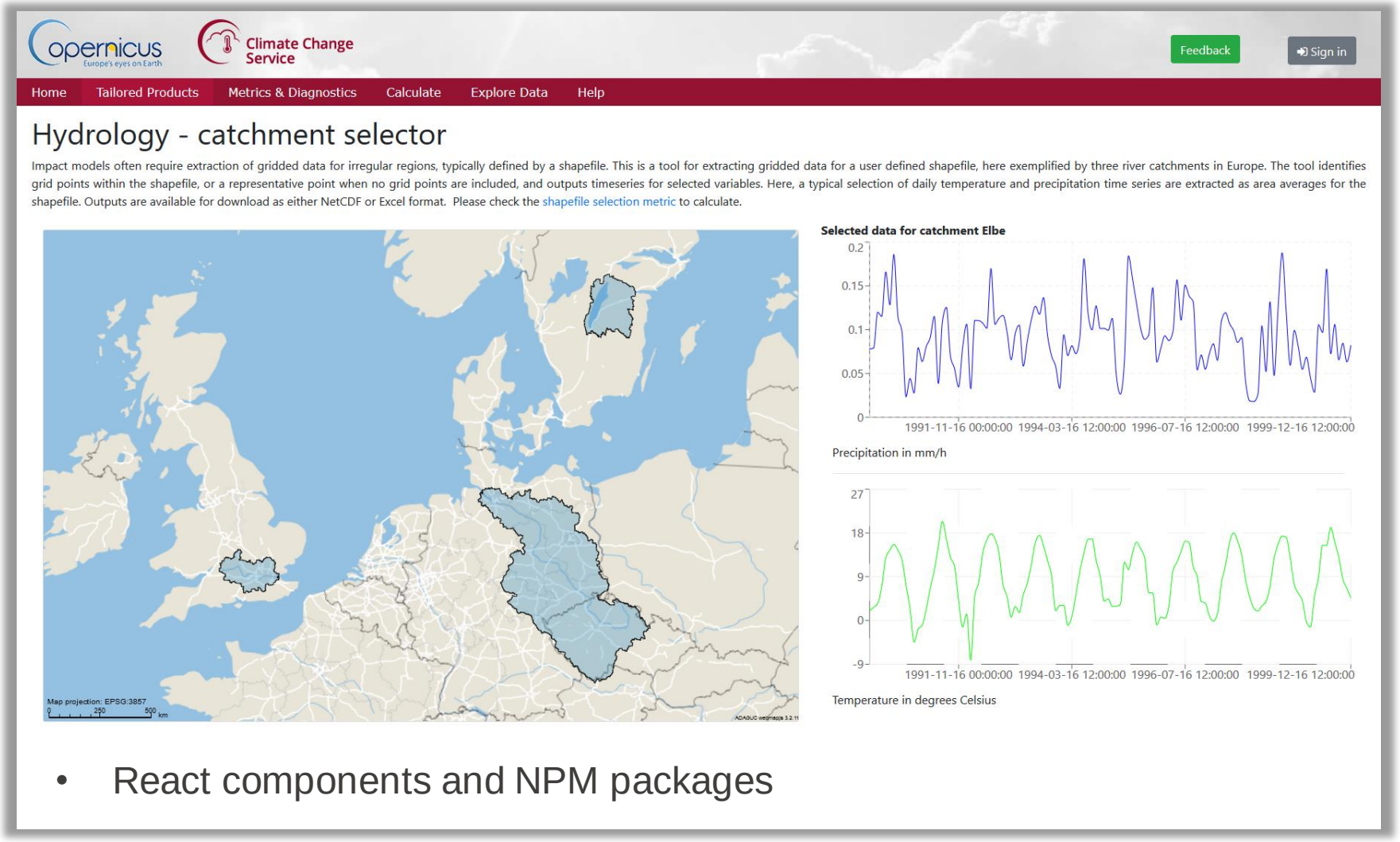

BY

Next steps ...

- ADAGUC-Services is a framework to build backend for portals
  - Java Spring based, offers open standards
  - Per user environment – including certificates
  - Sign in via OAuth2, in combination with CEDA's certificate service
  - Connects to ESGF – (*ESGF search, ESGF catalogs, ESGF data)*
  - Implements the "User basket" where personal files can be stored.
    - Share data, supports sharing of data,
    - Supports OpenDAP, Supports visualization via ADAGUC WMS
    - Files can be uploaded and used for own analysis.
  - Connects with remote WPS services, results are copied to local basket
    - Use of remote processing services like DKRZ's Birdhouse WPS framework



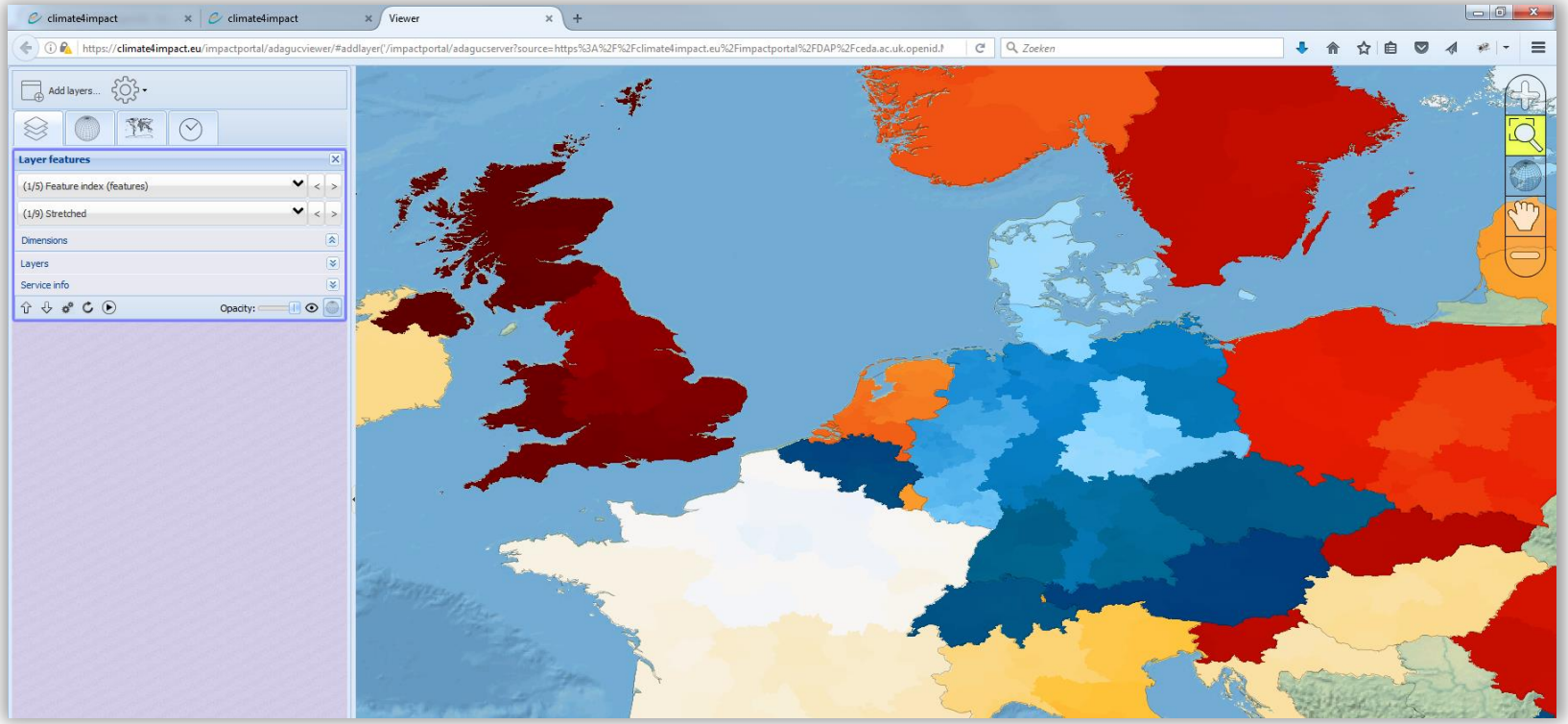
New frontend using ReactJS – based on previous work



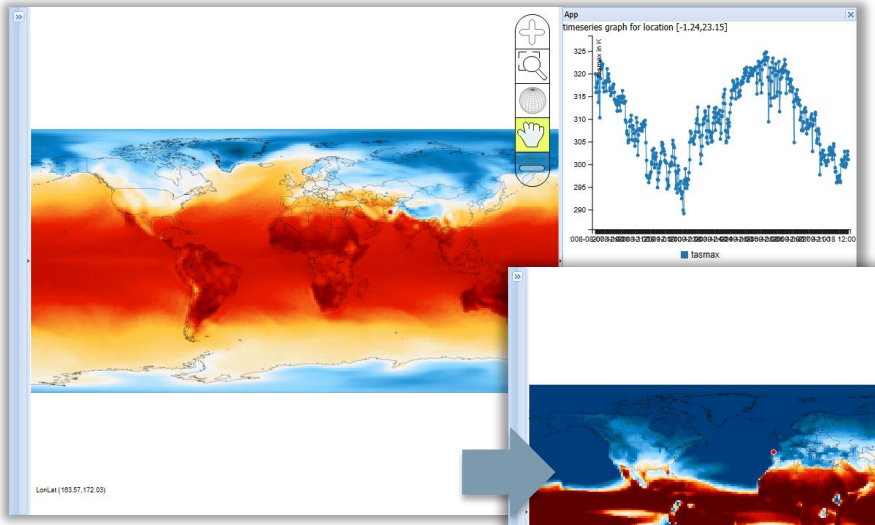
- React components and NPM packages

Use your own vector data defined in GeoJSON

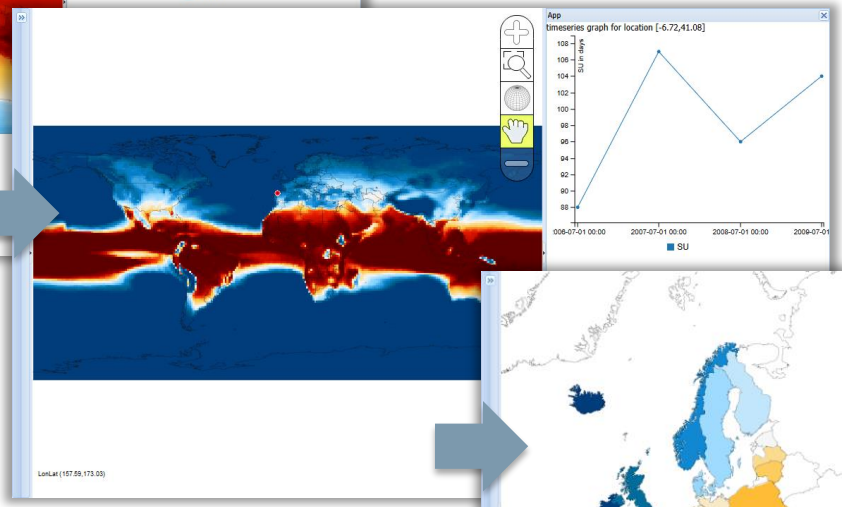
- ADAGUC-Server supports GeoJSON via WMS and WCS
- GeoJSON can be rasterized to any grid using the Web Coverage Service
- Allows for various operations like statistics per polygon, area extraction, etc..



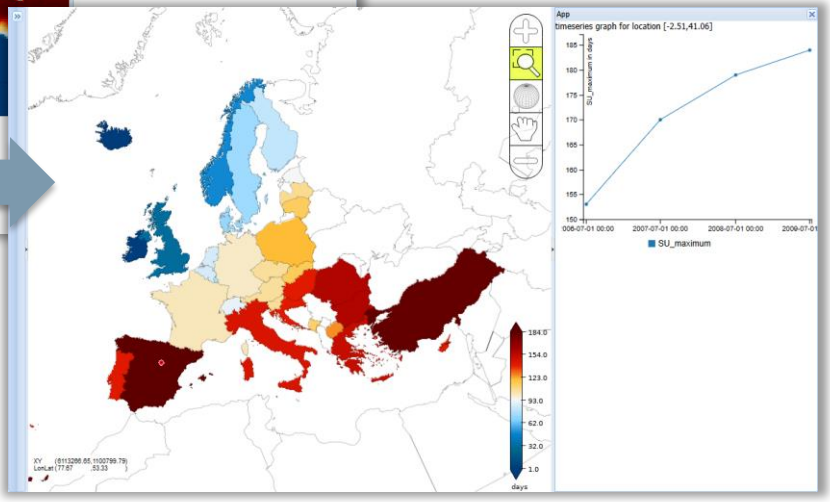
To summarize:



Daily Maximum temperature



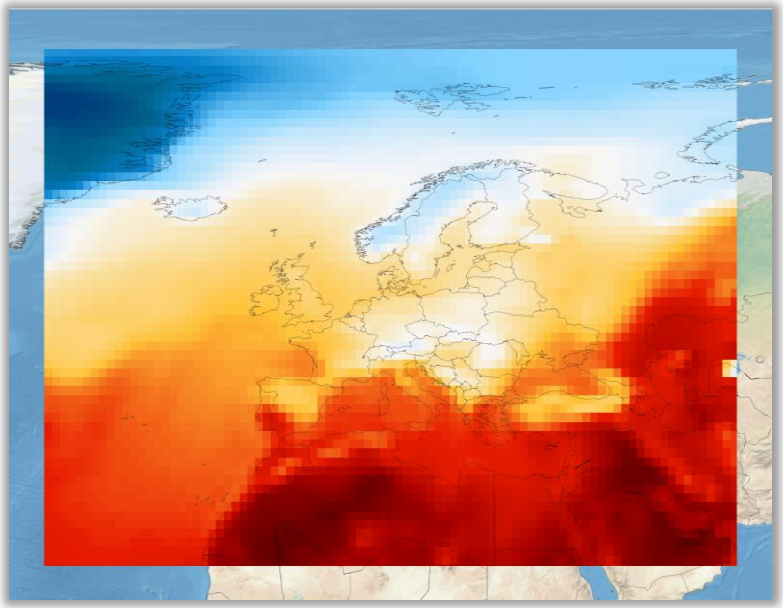
Summer days (T > 25)

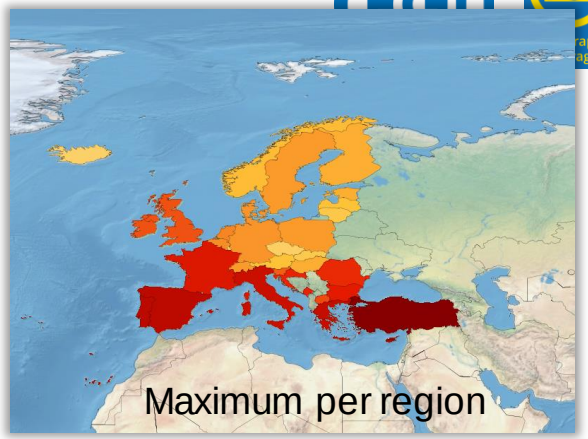


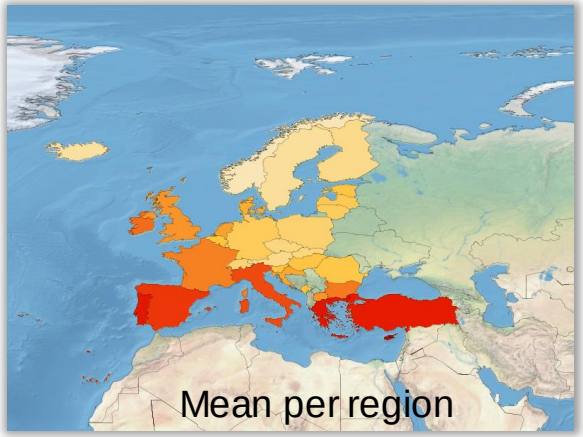
Maximum summer days per country

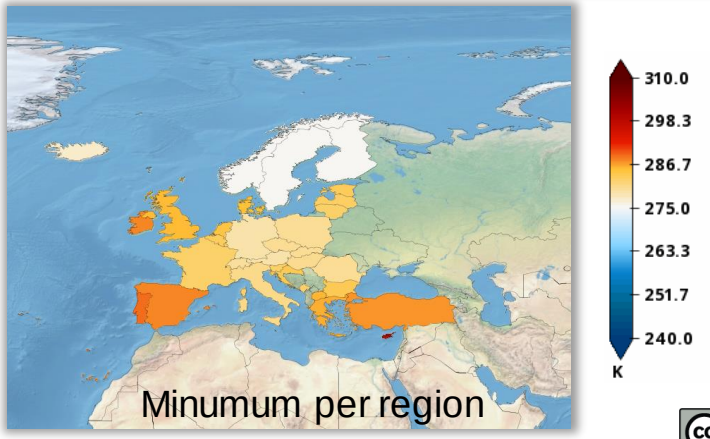
WPS for statistics per region

- Calculates statistics per region
- Mean, min, max, etc ...



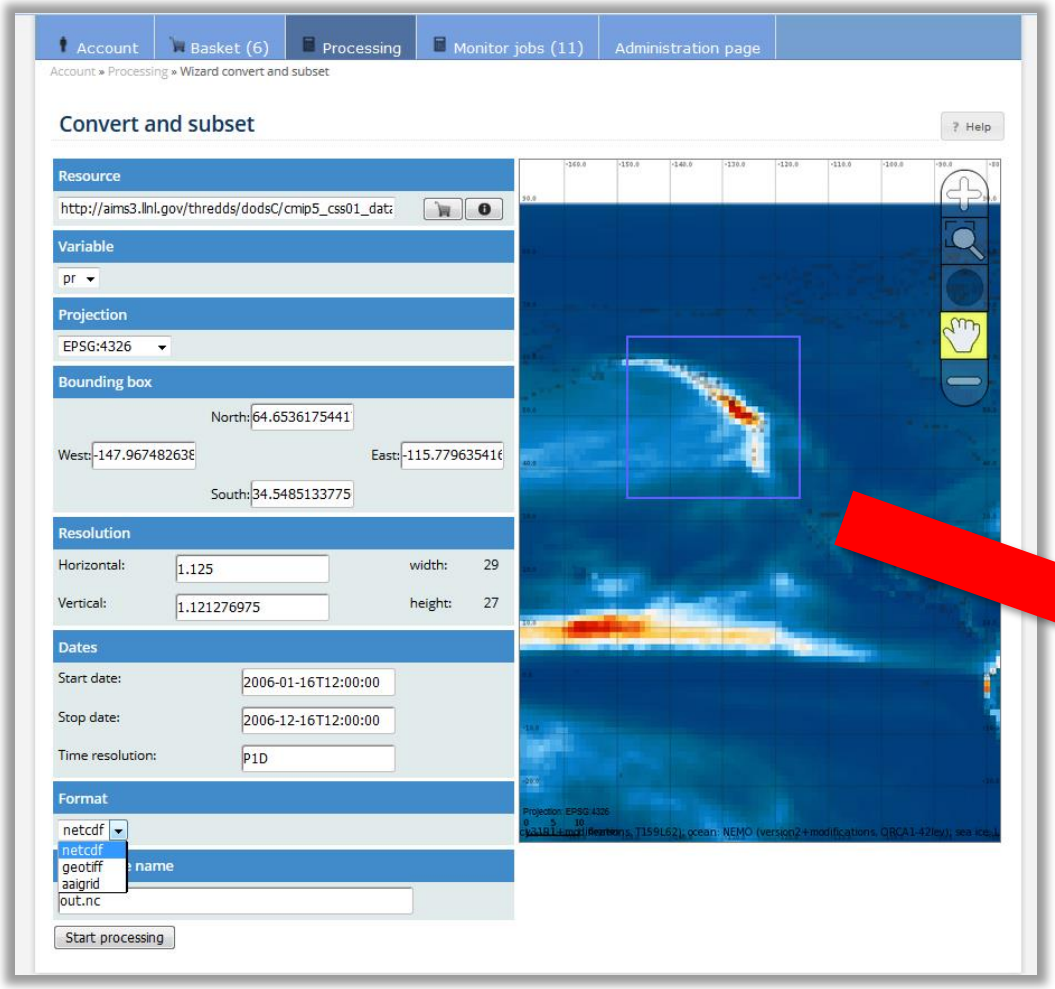




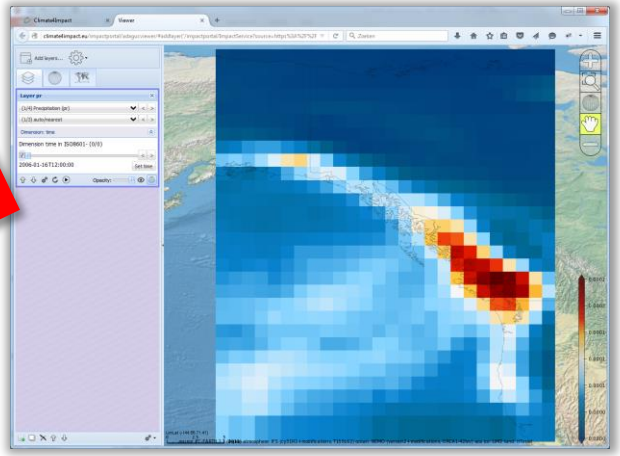


- Outputs CSV table and NetCDF files
- Based on NUTS regions in GeoJSON

Wizards for subsetting and regridding climate data

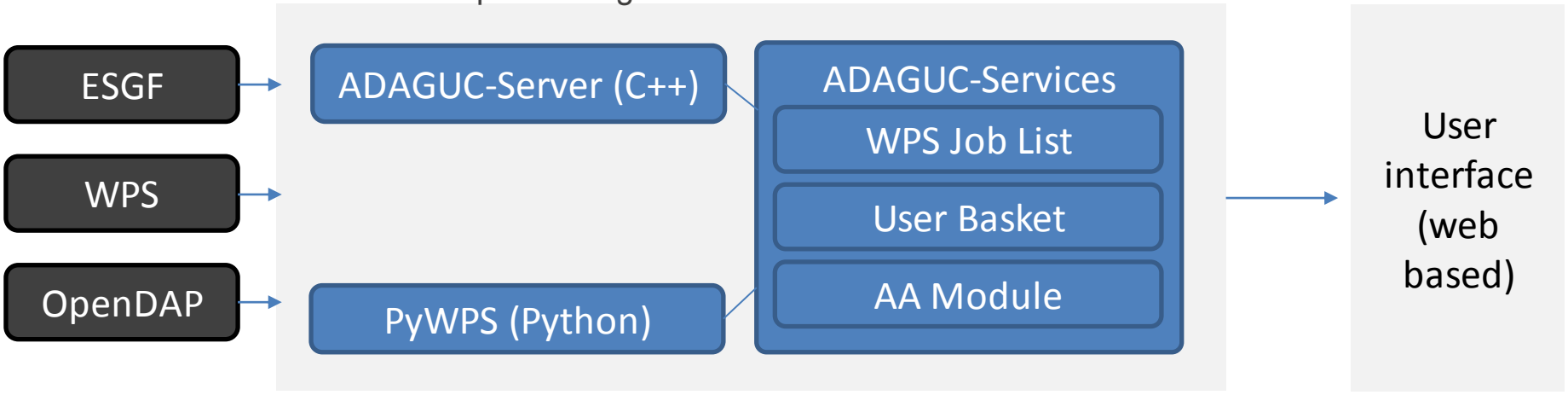


- Data reduction
- Subsetting
- GIS formats
- Reprojection
- Regridding



ADAGUC Services framework as backend

- Underlying technology (deployed at AWS and SurfSara HPC)
- ADAGUC-Services is a framework which wraps Adaguc-Server and PyWPS
  - Connects to ESGF – (*ESGF search, ESGF catalogs, ESGF data)*
  - Implements the “User basket” where personal files can be stored.
    - Supports visualization via ADAGUC WMS
    - Share data, supports sharing of data,
    - Supports OpenDAP.
    - Files can be uploaded.
  - Connects with remote WPS services, results are copied to local basket
    - Use of remote processing services like DKRZ’s Birdhouse WPS framework

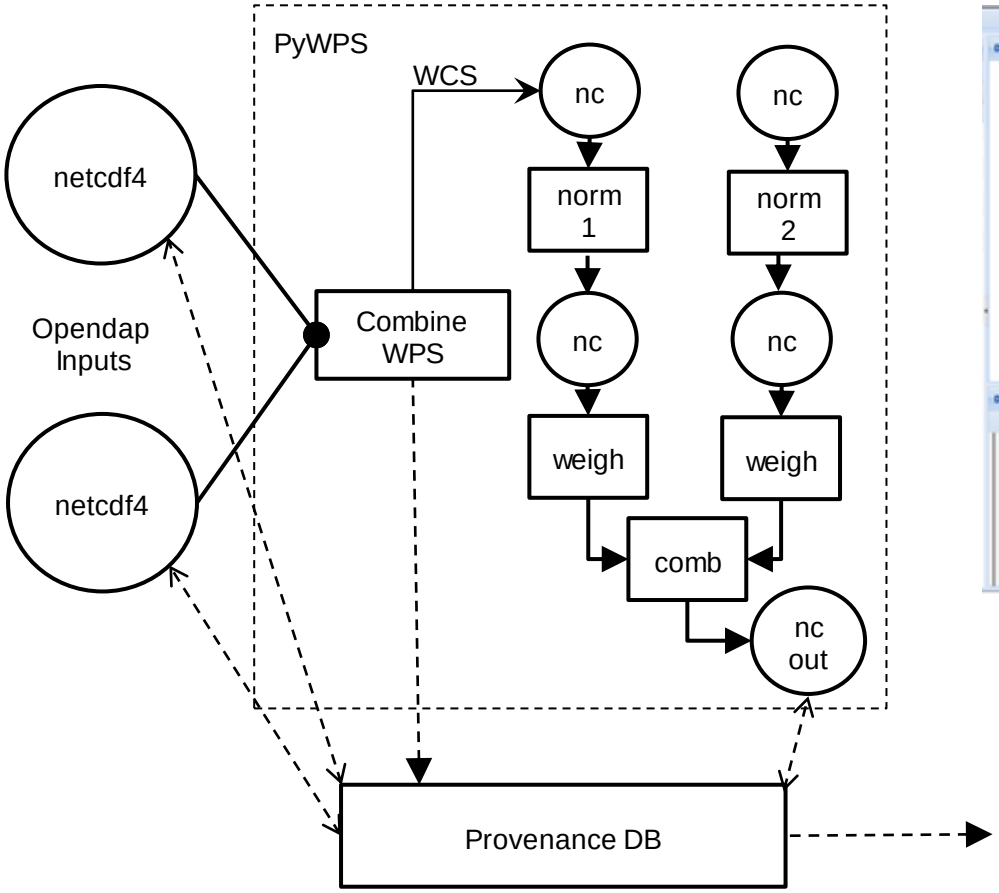


Use of adaguc-services backend framework

Workflows, Provenance, Tracability and Reproducability

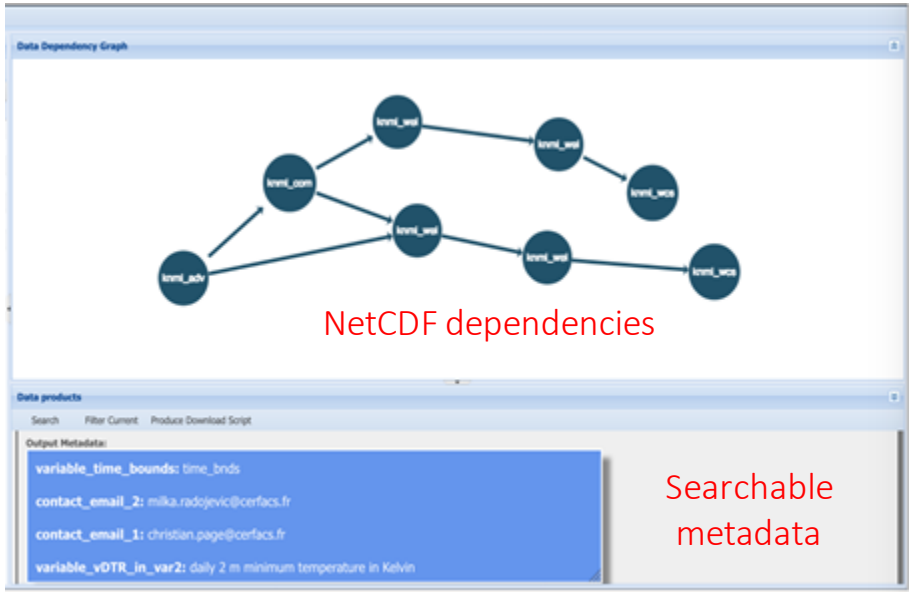
Combine WPS, five steps involved

- Provenance module: WPS_PROV
- Provenance metadata is stored in NetCDF
- W3C PROV-DM standard



Visual analytics techniques on provenance

- Highlighting data-reuse, even for cached data
- User interactions
- Exploitation of resources



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 824084

Backup slides