Climate4impact: Provenance in processing and new ADAGUC-services analytics

Maarten Plieger, Wim Som de Cerff, Alessandro Spinuso, Ernst de Vreede, Niels Drost, and Christian Pagé
Royal Netherlands Meteorological Institute (KNMI), R&D Observations and Data Technology, De Bilt, Netherlands
(maarten.plieger@knmi.nl)

The aim of climate4impact (C4I) is to enhance the use of research data and to support other climate portals. It has been developed within the European projects IS-ENES, IS-ENES2, CLIPC and C3S-34a lot 2 Magic. C4I utilizes the globally distributed ESGF webservices as data source, it uses ESGF search, OAuth2, Opendap and Thredds catalogs via certificate based authentication. C4I offers web interfaces for searching, visualizing, analyzing, processing and downloading (climate) datasets and is targeted to climate impact reseachers.

In the C3S-Magic project, the software and tools developed in C4I are re-used, extended and made open source. The analytics part is released under the ADAGUC-services flag, adding Web Processing Services to the already existing Web Map and Web Coverage services. Provenance integration is achieved using the W3C PROV standard for fully traceable provenance. The PROV document is stored in NetCDF files and can be visualized. The provenance module traces data usage statistics in a database which is interesting for data providers. Processing services include climate indicator calculations, country based statistics and polygon extraction by GeoJSON. The software is reusable, modular and packaged. Components are available via docker containers to enable easy re-use. Our tools are hosted on dockerhub and github for providing access to a wide audience.

ADAGUC-services also provides a personal basket where users can upload their own data and do research with the provided tools. The basket supports formats like NetCDF, GeoJSON and CSV. The basket has an access token mechanism to make data sharing and command line access to webservices easier. This makes it possible to provide services for third party portals, like the EU FP7 CLIPC portal. The CLIPC portal has for example a frontend more targeted to boundary workers than to researchers.

The processing capabilities, provided open source tools and provenance integration are detailed and can be demonstrated in this presentation.