



Dynamic Data Citation and Subset Creation for Climate Scenarios

Chris Schubert and Katharina Sack

Climate Change Centre Austria, Data Centre, Austria (chris.schubert@ccca.ac.at)

Citing datasets in an appropriate manner is recognised as good scientific practice. Dealing with the original source as part of the provenance within the data life cycle is an essential basis for research data to carry out verification, reproducibility and basic element against data manipulation.

The CCCA Data Centre has taken the development of a structured, web-based service that describes the relations between originals and derivatives as subsets, their versioning and an automated citation text. This application was formed as a RDA Data Citation working group pilot on NetCDF Climate Scenarios in cooperation with the research project e-infrastructures Austria.

Data content changes over a time line, will be corrected or replaced by new ones. To establish a reproducibly of data processes and their re- use, researchers need the ability to identify the exact version of a data set used.

If only subsets of a data set have been used, it is very time-consuming and complex to cite the original data and the versions based on it. The aim was to describe the processes for subset formation precisely, precisely and comprehensively in a web based framework.

The presentation gives a brief overview of the technical implementation, features developed for the creation of subsets, how to identify resources by storing queries and executing queries again as well the automatic creation of a landing page of subsets which provides a citation text as well inherited metadata in a dynamic manner.