

Assignment 3: Identifying Datasets

Table 1: Information about Data Repository	
Information Category	Information Identified
Name Of The Repository	World Data Center for Climate Additional name: DKRZ long term archive
URL Of The Repository	https://www.wdc-climate.de/ui/
Owner Of The Repository	German Climate Computing Centre (DKRZ), World Data System
Maintenance Authority	German Climate Computing Centre (DKRZ)
Funding Institution	German Climate Computing Centre
Time Coverage	2004 - present
Subject(S) Of The Repository	Natural sciences; Geosciences; Atmospheric science, Oceanography and Climate Research, Geology and Paleontology
Data Content	Climate science data and related data subjects such as earth system science, nature scientific data, marine sciences, terrestrial sciences, geosciences
Data Format(S)	NetCDF, GRIB, CSV, acii, GRIB1, IEG, netCDF.tar, LOCAL, tar.gz, GRIB1 zipped; recs separated, pdf, tar-File(s), IEEE, xml, GRIB2, GRIB1; records separated, zip-file, ASCII with prefix, other/diverse, netCDF.gz, netCDF.zip, ASCII.gz, netCDF.tar.gz, grib_rg, BIN.gz, INTEGER*2, ASCII records separated, GeoTIFF, ESRIarchinfoGrid.ascii.zip, xls, tar file zipped, tar of png, PNG, GRIB1;records separated;bz2-zip, grib_sz, PS, tar of grb, tar of jpeg, tar of ascii, DOC, GZ, tar of gif, zip of gif, HDF5, JPEG, XLSX, bin (WRF-WPS), ppt
Data Collection Method	Contribution from other environmental and climate data centers and scientist across the globe
Data Access Methods	Download/export, browsing, online analysis

Data Grouped Methods	Project (subject or type), aggregation, data type, organization, format, status, temporal coverage, code
Dataset List	Given the multiplicity of how the data can be categorized, there is no comprehensive list of all the datasets
Search Interface	Type-in keyword bar with filter options
Search Option	Keyword, code, longitude/latitude, temporal coverage

Table 2: Information about Dataset	
Information Category	Information Identified
Name Of The Dataset	Cops_suph_cradar_oview_p: overview pictures of 35GHz cloud radar data
URL Of The Dataset	https://www.wdc-climate.de/ui/entry?acronym=cops_suph_cradar_oview_p
Summary Of The Dataset	Pictures from cloud radar taken daily for precipitation prediction and analysis
Project	COPS (Convective and Orographically-induced Precipitation Study)
Original Language	German
Location	COPS-Supersite Hornisgrinde
Spatial Coverage	Longitude 8.1 to 8.3 Latitude 48.53 to 48.67 Altitude: 1050 m relative to NN to 15000 m relative to NN
Temporal Coverage	June 01, 2007 – August 31 - 2007
Data Catalog	World Data Centre for Climate
Dataset Size	26.51 MiB
Format	JPEG
Status	Completely archived
Creation Date	January 28, 2008
Access Type	Authorized Users only
Additional Project Information Link	https://www.wdc-climate.de/ui/entry?acronym=cops_suph_cradar_info_1

Summary: For this assignment, I chose the World Data Center for Climate Sciences as my repository focus and subsequently chose a dataset on cloud coverage for my dataset highlight. This repository holds well over two million datasets, with 51 data format types and 196 different projects. The global focus of this climate-focused informational hub is not short of worldly coverage; it is specific in subject matter that it will take on, relatedly only to climate change and effects, or other marine, atmospheric, and geoscience subjects. While I was initially intrigued by the notion of a climate science informational hub, it is clear to me that after exploring this resource, it is over-inundated with datasets and types and could benefit from a reconfiguration. WDCC holds 51 different types of file formats, some of which look to be duplicates or extremely slight variations of one another. Per the website's description and mission statement, it is open to any institution or scientist who would like to contribute new datasets or add to pre-existing projects. However, it does not seem to have utilized a more modernized or standardized system that could create more coherence or specification as to accessible data types. I argue it could also benefit its users and contributors by reformatting or updating older datasets to fit type configurations. Many datasets are also restricted for download, which does not make this interface as patron-friendly as it could be and is more likely designed only for institutions or scientists to have fuller access. My two suggestions for this repository would be to reorganize and redistribute: considering subject organization as an option for the search bar filter may help new users or non-science-facing patrons understand the potential of the repository as well as making either older data or some sections of data available to non-institutional users.