The Unidata Program Center develops software and advocates for standard approaches that make scientific data more accessible and interoperable between software packages. Unidata works with government agencies and standards organizations to promote wide use of techniques that make scientific data accessible now and in the future.

**NetCDF**

Unidata’s netCDF (network Common Data Form) is a freely distributed collection of data access libraries that provide a machine-independent data format that is self-describing, portable, scalable, appendable, sharable, and archivable – all important qualities for those who wish to create, access, and share array-oriented scientific data. NetCDF permits easy access to array-based, multi-dimensional datasets, a task that can be difficult when using other common storage schemes. NetCDF has been adopted widely by the atmospheric sciences community, defined as a standard by the Open Geospatial Consortium (OGC), and is especially popular among climate and ocean modelers. For example, model output datasets for the Fifth Assessment Report of the Intergovernmental Panel on Climate Change must be submitted in netCDF format, using the associated Climate and Forecast (CF) metadata conventions. The resulting large base of netCDF users and data has led to support for the format in more than 80 open source packages and many commercial applications including ArcGIS, MATLAB, and IDL.

**Common Data Model**

Unidata’s Common Data Model (CDM) provides an interface for reading and writing files in netCDF and a variety of other scientific data formats. The CDM uses metadata to provide a high-level interface to geoscience-specific features of datasets, including geolocation and data subsetting in coordinate space. Unidata’s THREDDS Data Server (TDS) builds on the CDM to allow for browsing and accessing collections of scientific data via electronic networks. Data published on a TDS are accessible through a variety of remote data access protocols including OPeNDAP, OGC Web Map Service (WMS) and Web Coverage Service (WCS), NetCDF Subset Service (NCSS), and HTTP.

The CDM and TDS are widely used in the United States (by NOAA, USGS, NASA, and the Earth System Grid, for example) and internationally, and are part of the deep infrastructure on which next generation capabilities are being built by other organizations. Additionally, many other tools build on the CDM (NOAA’s ERDDAP, NASA’s Panoply, and CMAS’ VERDI, are examples.)
Climate and Forecast Metadata Conventions

The conventions for climate and forecast (CF) metadata are designed to promote the processing and sharing of files created with the NetCDF API. The conventions define metadata that provide a definitive description of what the data in each variable represents, and the spatial and temporal properties of the data. This enables users of data from different sources to decide which quantities are comparable, and facilitates building applications with powerful extraction, regridding, and display capabilities.

The CF conventions are increasingly gaining acceptance and have been adopted by a number of projects and groups – including the Coupled Model Intercomparison Project Phase 5 (CMIP5) – as a primary standard.

Rosetta

The Rosetta project at the UPC is an effort to improve the quality and accessibility of observational data sets collected via datalogging equipment. The initial goal of Rosetta is to transform unstructured ASCII data files of the type commonly generated by datalogging equipment into the netCDF format, while minimizing disruption to existing scientific workflows.