Climate science support from the Research Data Archive at NCAR

Doug Schuster[†]; Steve Worley; Joey Comeaux; Tom Cram; Bob Dattore; Zaihua Ji; Chi-Fan Shih; Dave Stepaniak; Gregg Walters
[†] National Center for Atmospheric Research, USA
Leading author: schuster@ucar.edu

Improving discovery and access to high quality climate data is critical to advancing our understanding of the Earth's climate system and how it changes over time. The Research Data Archive (RDA), managed by the Computational and Information Systems Laboratory at NCAR, is a historical and growing data repository designed to meet the needs of the climate research community. The core datasets in the RDA are most basically characterized as collections of meteorological and oceanographic in situ observations collected from 1662 to the present, and a wide variety of analyses and re-analyses derived from them. Stewardship work on this archive began four decades ago with a purpose to serve research at NCAR. Since then, the RDA has grown rapidly and because of open access principles and system design, it provides services to users worldwide. During this past year more than 7000 unique users have extracted 260 TB of data from the RDA. This poster will describe core dataset categories that support climate research and the faceted and broad textual searches based on rich metadata that enable data discovery. The several data access options and ways that users can customize the data they receive will be shown. Over 200 TB (125 datasets) of the most sought after data are stored on disk and are directly available through the web server. Much of the remaining RDA, that totals over 600 TB and 600 datasets, can be requested and automatically restaged to disk from the NCAR mass storage systems. User selections of space, time, and parameter on web interfaces are used to create file lists and trigger data extractions from many TB size archives. Downloading options can be simple, directly through a browser (suitable for limited amounts of data), or accomplished by running scripts on your local machine that are prepared by the RDA system. A growing service feature is user determined output data format. The RDA has a supporting data-consulting arm provided by the staff that all have science educations and experience in research.