Unidata ([http://www.unidata.ucar.edu](http://www.unidata.ucar.edu/)) is a community data facility for the atmospheric and related sciences, established in 1984 by U.S. universities with sponsorship from the National Science Foundation (NSF). The Unidata Program Center (UPC), the program office for Unidata and the nexus of activities related to Unidata’s mission, is managed by the University Corporation for Atmospheric Research (UCAR), a consortium of over 109 member universities and academic affiliates providing science in service to society.

Unidata exists to engage and serve researchers and educators dedicated to advancing the frontiers of Earth System science. The program’s aim is to help transform the conduct of research and education in atmospheric and related sciences by providing well-integrated, end-to-end data services and tools that address many aspects of the scientific data lifecycle, from locating and retrieving useful data, through the process of analyzing and visualizing data either locally or remotely, to curating and sharing the results.

Specifically, the UPC:

* Acquires, distributes, and provides remote access to real-time meteorological data.
* Develops software for accessing, managing, analyzing, visualizing, and effectively using geoscience data.
* Provides comprehensive training and support to users of its products and services.
* In partnership with others, facilitates the advancement of tools, standards and conventions.
* Provides leadership in cyberinfrastructure and fosters adoption of new tools and techniques.
* Assesses and responds to community needs, fostering community interaction and engagement to promote sharing of data, tools, and ideas.
* Advocates on behalf of the community on data matters, negotiating data and software agreements.
* Grants equipment awards to universities to enable and enhance participation in Unidata.

Unidata is governed by its community. Representatives from universities populate standing and *ad hoc* committees that set policies for the program, provide first-hand feedback from users of program software and services, and offer guidance on individual projects

While Unidata’s primary mission of serving universities engaged in atmospheric science education and research has remained unchanged through the years, the evolution and broad usefulness of its products and services have greatly enlarged its initial user base. Today, the Unidata community includes users from all sectors in over 200 countries, including nearly 2500 academic institutions and more than 80 research labs. Simultaneously, Unidata’s activities and responsibilities have also grown as community needs have evolved. Despite the growth in users and enhanced scope of its activities, according to a 2010 survey conducted by the Unidata Users Committee, 97% of the respondents indicated that they were either satisfied or highly satisfied with Unidata’s overall service to the community.

In the following sections we highlight some key quantitative and qualitative metrics that are used to gauge Unidata’s success. These indicators offer a peek at Unidata’s impact and how its cyberinfrastructure plays an irreplaceable role in advancing research, education, and outreach goals of its community. It should be noted that the UPC provides many of these metrics to its governing committees as part of its regular status reports.

## Data services

|  |  |  |  |
| --- | --- | --- | --- |
| Table 1: Growth of IDD data volume | 2008 | 2013 | 2016 |
| Volume of data pushed to IDD sites | 2.7 TB/day | 13 TB/day | 18.8 TB/day |

Delivery of geoscience data to universities in near real time via the IDD system is at the core of Unidata’s mission and is extremely important to our university community. Table 1 gives an idea of the explosive growth in the volume of data delivered via the IDD over the past decade.

While the IDD uses a “push” mechanism to deliver data automatically as it becomes available, Unidata’s remote data access mechanisms (including THREDDS Data Servers, ADDE servers, RAMADDA servers, and EDEX servers) also provide roughly **670 GB/day** to community members.

## Software and support

|  |  |
| --- | --- |
| Table 2: Software Package Downloads | 2012-2016 |
| AWIPS | 7700 |
| GEMPAK | 8400 |
| IDV | 43600 |
| LDM | 10100 |
| McIDAS | 300 |
| netCDF-C Libraries\* | 566800 |
| netCDF-Java Libraries (Common Data Model) | 41900 |
| TDS | 9000 |
| UDUNITS | 22700 |

\* UPC source code downloads only. This number does not include downloads from source code repositories at the University of Kyoto or on the Github site, or binary distributions available via package managers on UNIX-like systems.

Unidata community members rely on the UPC to provide access to a variety of software packages for data transport, management, analysis, and visualization.  Table 2 shows how many community members have downloaded the software packages that the UPC develops and supports over the past five years.

In addition to providing the software for download, UPC developers also provide the community with direct technical support via electronic mail. The support system is heavily used, with **more than 21000 support queries handled by UPC staff in the past five years**.

## **Appendix: A description of the key products/services**

#### Data Distribution

The UPC coordinates the Internet Data Distribution system (IDD), in which hundreds of universities cooperate to disseminate near real-time earth observations via the Internet. While the “push” data services provided by the IDD system are the backbone of Unidata’s data distribution services, the UPC also provides on-demand “pull” data services via THREDDS, ADDE, and RAMADDA data servers.
The UPC’s data servers are not classified as “operational” resources, but they nonetheless have a 99.96% uptime record and are used heavily by educational sites that lack the resources to store IDD-provided data locally, or to operate their own data servers (see Table 1). UPC’s servers are housed in a UCAR co-location computer facility for reliability, and share UCAR’s Internet2/National Lambda Rail connectivity, which provides access to ample bandwidth for Unidata’s needs.

## **Software**

A variety of software packages are developed, maintained, and supported by the UPC:

#### 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, and is especially popular among climate and ocean modelers. For example, model output datasets for the Sixth 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 MATLAB and IDL.

#### Common Data Model & THREDDS Data Server

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.

#### Integrated Data Viewer

Unidata’s Integrated Data Viewer (IDV) is a 3D geoscience visualization and analysis tool that gives users the ability to view and analyze a rich set of geoscience data in an integrated fashion. The IDV brings together the ability to display and analyze satellite imagery, gridded data (such as numerical weather prediction model output), surface observations (METARs), upper air soundings, NWS NEXRAD Level II and Level III RADAR data, NOAA National Profiler Network data, and GIS data, all within a unified interface. The IDV integrates tightly with common scientific data servers (including Unidata’s TDS) to provide easy access to many real-time and archive datasets. It also provides collaborative features that enable users to easily share their own data holdings and analysis products with others.

#### AWIPS II & GEMPAK

AWIPS II is a weather forecasting, display, and analysis package currently being developed by the NWS and NCEP. Because many university meteorology programs are eager to use the same tools used by NWS forecasters, Unidata community interest in AWIPS II is high. UPC staff have worked closely with NCEP staff during AWIPS II development in order to devise a way to make it available to the university community.

NCEP has stated that GEMPAK applications will be migrated from GEMPAK/NAWIPS into AWIPS II for the National Centers. The UPC will likewise facilitate a migration from GEMPAK/NAWIPS to AWIPS II for the university community.

#### 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.

#### Local Data Manager

The Unidata Local Data Manager (LDM) system includes network client and server programs designed for event-driven data distribution. It is the fundamental component of the IDD system. The LDM is used by hundreds of sites worldwide, and is integrated into the National Weather Service’s AWIPS II package.

#### McIDAS

The Man-computer Interactive Data Access System (McIDAS) is a large, research-quality suite of applications used for decoding, analyzing, and displaying meteorological data. The older McIDAS-X system, developed by the University of Wisconsin’s Space Science Engineering Center and supported by Unidata, is gradually being replaced by the IDV and by McIDAS-V (which is based on the IDV).

#### UDUNITS

Unidata’s UDUNITS supports conversion of unit specifications between formatted and binary forms, arithmetic manipulation of units, and conversion of values between compatible scales of measurement.

#### RAMADDA

TheRepository for Archiving, Managing and Accessing Diverse Data (RAMADDA) is a vibrant and growing technology initially developed by Unidata and now managed and developed as an open source project. Unidata integrates RAMADDA functionality into the IDV, provides training and support, and contributes code to the project. In addition, Unidata makes extensive use of RAMADDA to support community and collaborative projects, and actively facilitates its deployment in the university community.