



# EOSDIS Earth Observing System Data and Information System

Published on EOSDIS - Earth Data Website (<https://earthdata.nasa.gov>)

[Home](#) > [User Resources](#) > [Sensing Our Planet](#) > Overcoming Data Incompatibility

## Overcoming Data Incompatibility [1]

by Rachel Hauser  
Published in 1998

Scientists wish handling science data were as easy as surfing the Internet, so they could focus on algorithms and analyses, rather than on the bookkeeping and format conversions that take so much of their time. Getting data is a chore. Not only must scientists sift through terabytes of data to obtain a few essential bytes, but after sifting they must often convert these data to a format their software can read.

What if you could select Earth science data as easily as you could browse the Web?

- [About PO.DAAC](#) [2]

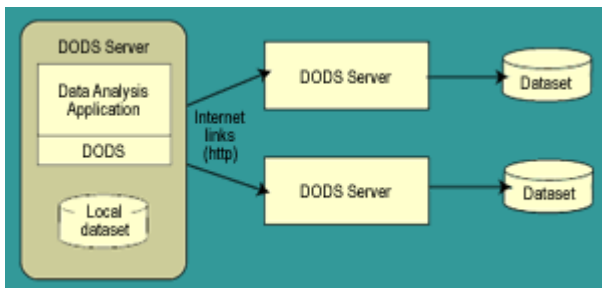
But what if you could sift through the huge amounts of Earth science data available and select whatever small pieces you need directly into your computer program for processing, as easily as you could browse the web? What if all that work were handled behind-scenes, as adroitly as your web browser seeks and presents pages without troubling you for help in orchestrating the nuances -- even though the details (text in its myriad headlines, fonts, tables, columns; imagery in different file formats, sounds, animations, and interactive virtual-reality animations) are seriously complex? Would we be in Kansas anymore if science data files, no matter the complexity and no matter the provenance, were each handled swiftly and invisibly by a program in your computer -- all with no help from you?

Feedback

"Instead, data are typically written in different formats and subsetted according to data providers' whims," said Peter Cornillon, oceanographer at the University of Rhode Island (URI). "In one instance, I downloaded 300 megabytes of data to get the 13 megabytes I needed. Then, I had to acquire a program capable of interpreting the data, after which I had to transfer the data to my own software application."

In an effort to overcome such problems and enhance his own research collaboration capabilities, Cornillon, James Gallagher of URI, and Glenn Flierl, professor of oceanography at MIT, developed the Distributed Oceanographic Data System (DODS) to provide access to each other's oceanographic data. DODS permitted them effortless data access and interpretation despite differences in software applications. It also facilitated extraction of data subsets, such as air temperature and wind speed on a particular day.

"We wanted to move data easily from one place on a network to another without having to worry about what format the data were in," said Cornillon. "DODS allows you to open your software application, such as Matlab or IDL, and access the data you need using the World Wide Web's client/server technology."



DODS architecture (Image courtesy of DODS)

Realizing that colleagues experienced similar data access problems and that DODS might prove useful to a wider audience, Cornillon and Flierl obtained funding from NASA and NOAA to adapt DODS' capabilities to solve data dilemmas for other scientists requiring oceanographic and meteorological data. They consulted outside academic and government laboratories to determine system requirements, and developed the existing web-based design.

Finding data on DODS is similar to conducting an Internet search. Users searching out variables from a particular data set submit a query to a DODS server. The DODS client allows the user to create a URL containing the server's web address as well as any relevant data constraints. Because a DODS request inherently contains a lot of information packed into one URL string, writing these strings quickly becomes complicated. Graphical user interfaces facilitate the request process.

"A request can be made from a researcher's application program, which sends a URL describing the data format, server name, data location, file name, as well as a constraint expression such as a variable name, and the data range," Cornillon said. "The DODS server sorts out the request and returns all of the relevant data."

"Format interoperability is important -- DODS takes care of translating between the data storage format and the chosen software application, allowing the researcher to employ existing application software," said Elaine Dobinson, deputy manager of the Jet Propulsion Lab's (JPL) Physical Oceanography Distributed Active Archive Center (PO.DAAC).

Dobinson realized that DODS could allow researchers to gain access to the vast stores of NASA's Earth Observing System (EOS) data without having to understand Hierarchical Data Format (HDF), EOS' chosen data storage format. So, in 1995 Dobinson suggested that JPL create an HDF DODS server.

"Building the HDF server for EOS data was our first DODS-related activity. We found that it was difficult to search DODS data sets using spatial and temporal constraints so we suggested building a graphical user interface that allowed users to search the DODS system in ways that were familiar to them," she said.

The JPL team created an interface to serve as a prototype for the calibration of the NASA Scatterometer (NSCAT) instrument, said Dobinson. "Our idea was to provide an interface for the team to use ground truth data sets to collocate and verify NSCAT data. We took a few sample data sets, put them in DODS, and created an interface capable of plotting the spatial and temporal coverage of the selected data so that the user could pick out a data subset by merely pushing a button," she said.

In addition to JPL's HDF interface, a Matlab interface built by the DODS core group provides users with a simple list of variables and the 30 most requested data sets.

"Right now over 300 data sets are available, consisting mostly of meteorological and oceanographic data. Many of those are very small, but we add about one new data set per week," said Cornillon.

The system works well now, but it will soon become unwieldy due to sheer data set numbers, said Dobinson. "A better search interface needs to be built," she said.

System constraints still exist. For example, DODS currently works only on UNIX machines (although a DODS port to Windows NT is in progress at JPL) and keywords describing the data variables are not available for every data set, but each problem encountered and solved has increased the functionality of the system.

"There are still some bumps in the road, users will encounter some problems, but the system has been built and the programs that move the data around have been in place for two to three years and are robust. We're working on the bells and whistles now," said Cornillon. "I had this dream of what this system would be like but I have to admit, I never thought we'd get there."

## Related Link(s):

- [Distributed Oceanographic Data System](#) [3]

**Source URL:** <https://earthdata.nasa.gov/featured-stories/featured-research/overcoming-data-incompatibility>

### Links:

[1] <https://earthdata.nasa.gov/featured-stories/featured-research/overcoming-data-incompatibility>

[2] <https://earthdata.nasa.gov/data/data-centers/po-daac>

[3] <http://www.unidata.ucar.edu/packages/dods/>