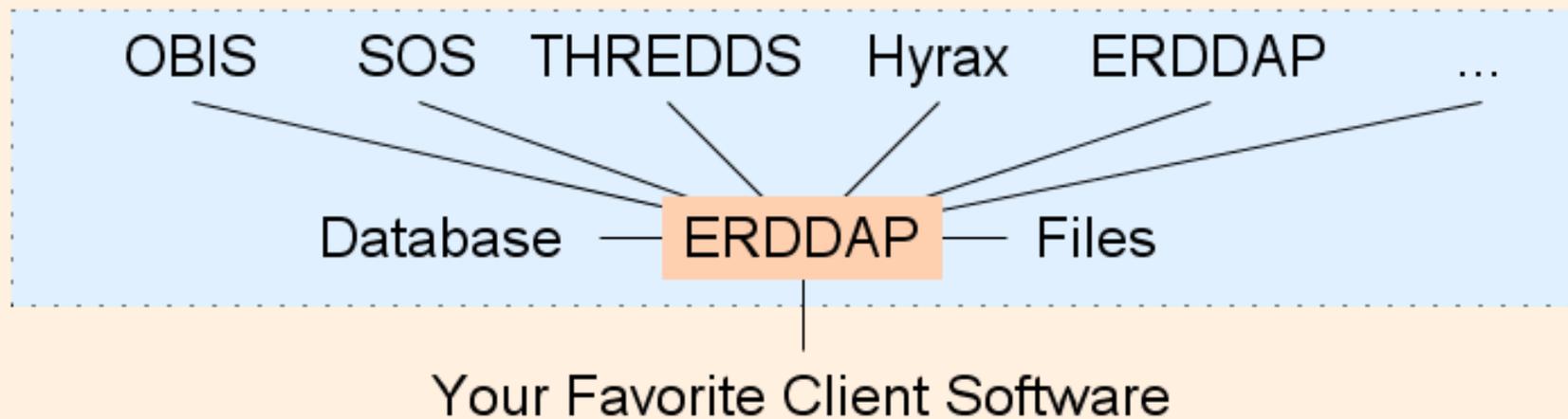


ERDDAP

Environmental Research Division's Data Access Program

ERDDAP is a data server designed to help you get data from any data server into your favorite client software.



Try it: coastwatch.pfel.noaa.gov/erddap
Bob Simons <bob.simons@noaa.gov>
NOAA NMFS SWFSC ERD

Downloading data is MUCH harder than it should be. ERDDAP tries to solve these problems:

↙ "Use case"

When a scientist or a computer program looks on the internet for data (e.g., sea surface temperature data), there are problems ...

1. **Search** - The datasets of interest are hard to find because they are at many different web sites.
2. **Requests** - Each site requires a different protocol to request data. (XML, SOAP+XML, OPeNDAP, WCS, WFS, SOS, an HTML form, ... !)
3. **Responses** - Each site returns the data in a different format (XML, SOAP+XML, DAP binary, ASCII text, HDF 4, HDF 5, NetCDF, ... !)
and it isn't the common file format that you want.
(.html table, ESRI .asc, .kml, .mat, .nc, .csv, .tsv, .json, .xhtml, ... !)
4. **Incompatible Time Formats** make it hard to compare data from different datasets. ("Jan 2, 1985", "02-JAN-1985", "1/2/85", "2/1/85", "1985-01-02", Year + DayOfYear, "seconds since 1988-01-01", "days since has 1-1-1", ... !)
5. **Lack of Metadata** (information about the data) makes it hard to understand the data.

Problem #1: Search

Datasets of interest are hard to find because they are at many different web sites.

ERDDAP's Solution

- ERDDAP is a middleman between you and remote data servers. This lets ERDDAP act as if it has lots of datasets. A search for datasets of interest searches all of those datasets.
- ERDDAP offers Google-like full-text search
For example: sea surface temperature
Also available via an OpenSearch service.
- ERDDAP offers category (faceted) search
For example: standard_name = sea_surface_temperature
The ERDDAP administrator specifies the categories.

Problem #2: Requests

Each data server requires a different protocol to request data.

(HTTP GET, XML, SOAP+XML, DAP, WCS, WFS, SOS, an HTML form, ... !)

Problem #3: Responses

Each data server returns the data in a different format

(XML, SOAP+XML, DAP binary, ASCII text, HDF 4, HDF 5, NetCDF, ... !)

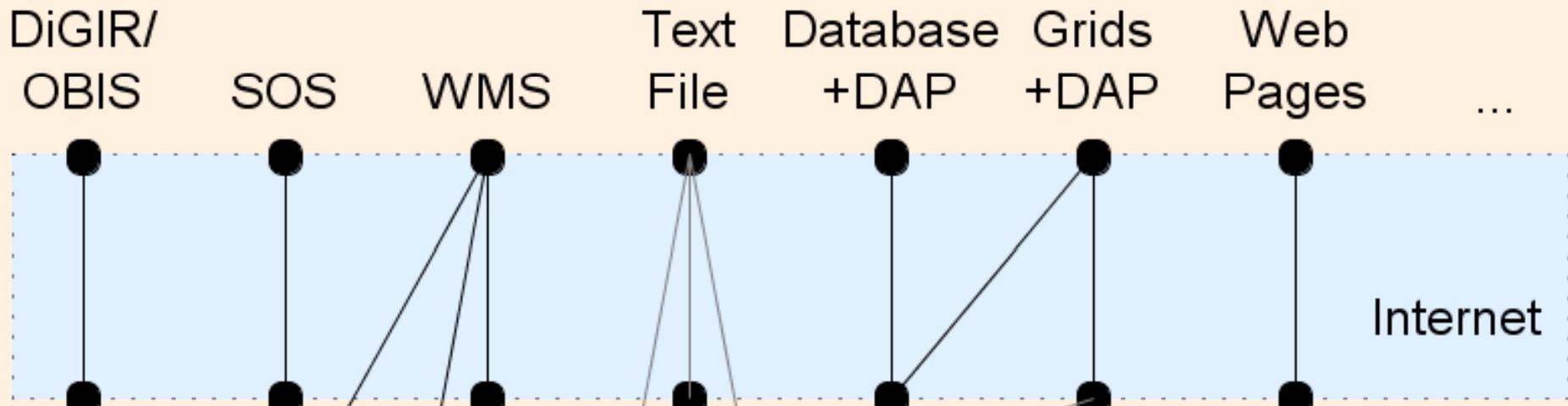
and it isn't the common file format that you want.

(.html table, ESRI .asc, .kml, .mat, .nc, .csv, .tsv, .json, .xhtml, ... !)

Getting Data without ERDDAP

Each type of data server is great with its own client(s).

Internet Data Server Types



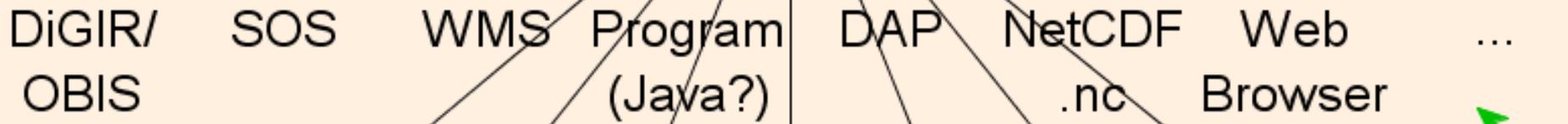
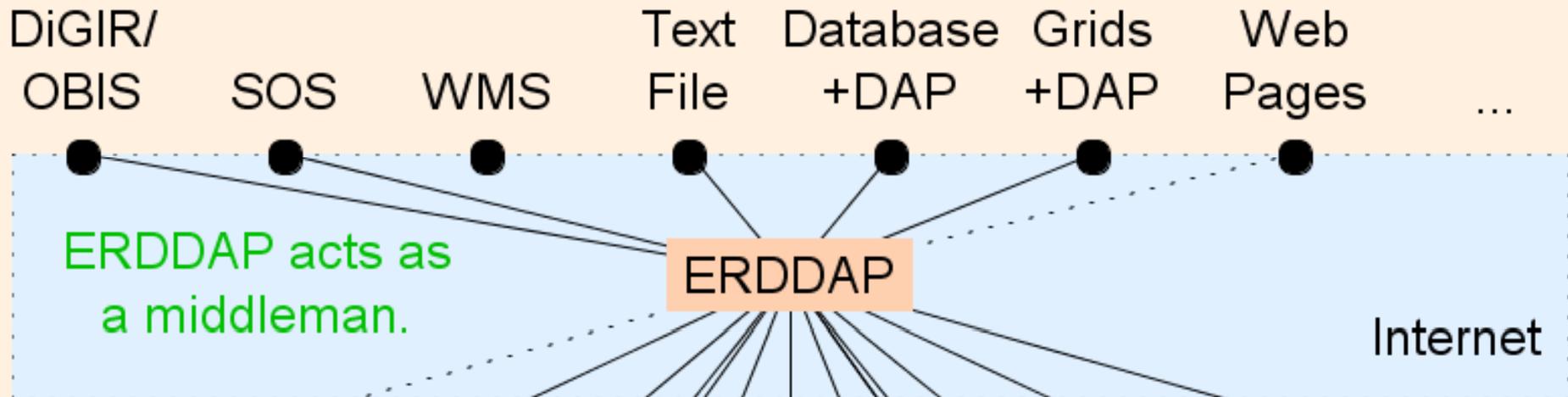
Internet Data Client Types



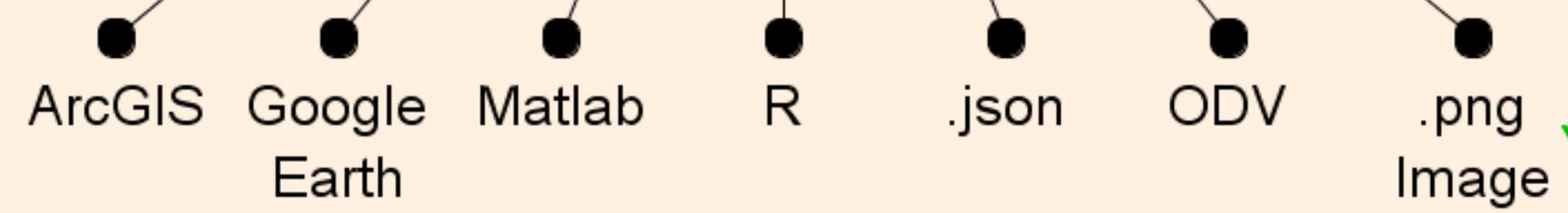
But there's no easy way to get data into other programs or file types.

Getting Data with ERDDAP

Internet Data Server Types



You can use your favorite client to get data from many sources.



You can get data into many common programs and file types.

ERDDAP's Solution

- ERDDAP is a middleman between you and remote data servers. You get a consistent way to request data. You get more options for the format of the response.
- Internally, ERDDAP uses just two simple data models
 - Multi-dimensional grids
 - Database-like tables

That allows ERDDAP to be an OPeNDAP server

(.dds, .das, .html, .asc, .dods)

with additional file extensions

(.csv, .esriAsc, .json, .mat, .nc, .odvTxt, ... for data

and .graph, .kml, .png, .pdf, for graphics)

and a WMS server.

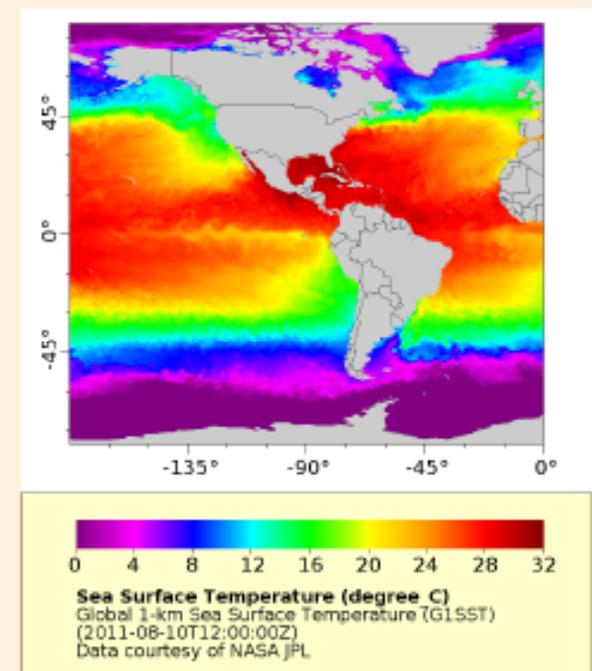
Griddap - grid data and graphs via DAP hyperslab requests

You can request data via a specially formed URL.

- The URL specifies everything: dataset, response file type, subset:

[http://coastwatch.pfel.noaa.gov/erddap/griddap/jplG1SST.html?sst\[\(2011-08-10T12:00:00\)\]\[\(-80\):100:\(80\)\]\[\(-180\):100:\(0\)\]](http://coastwatch.pfel.noaa.gov/erddap/griddap/jplG1SST.html?sst[(2011-08-10T12:00:00)][(-80):100:(80)][(-180):100:(0)])

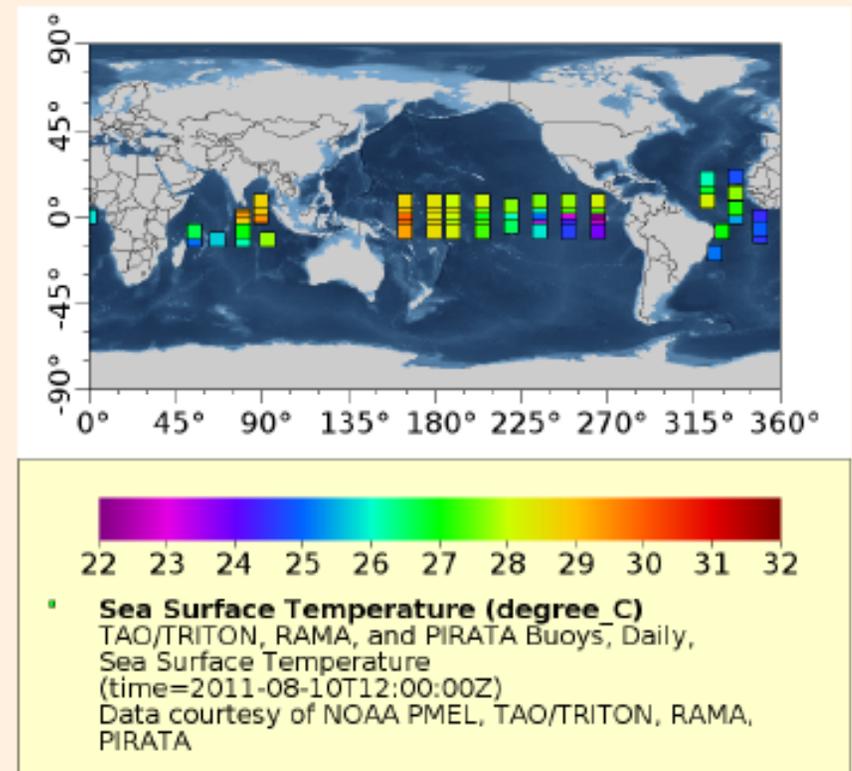
- Special file types: .html (Data Access Form), .graph (graphical form)
- Data file types: .asc, .csv, .das, .dds, .dods, .esriAscii, .htmlTable, .json, .mat, .nc, .ncHeader, .tsv, .xhtml.
- Image file types: .geotif, .kml, .smallPdf, .pdf, .largePdf, .smallPng, .png, .largePng, .transparentPng.
- If time is `last`, you will get the latest data.



Tabledap - tabular data and graphs via DAP constraint requests

You can request data via a specially formed URL.

- The URL specifies everything: dataset, response file type, subset:
http://coastwatch.pfel.noaa.gov/erddap/tabledap/pmeltaoDySst.html?longitude,latitude,T_25,time&time=2011-08-10T12:00:00
- Special file types: .html (Data Access Form), .graph (graphical form)
- Data file types: .asc, .csv, .das, .dds, .dods, .geoJson, .htmlTable, .json, .mat, .nc, .ncCF, .ncHeader, .tsv, .xhtml.
- Image file types: .kml, .smallPdf, .pdf, .largePdf, .smallPng, .png, .largePng.



ERDDAP: Web Applications Built on Web Services

Web Application - a web page with a form

Great for humans with a browser, but can't be automated!

(Imagine making 1000 requests by hand!)

Web Service - a request URL generates a response (REST)

Great for computer programs and scripts (easy to automate!)

- Standard request format (one URL specifies everything)
- Standard response format(s)

So ERDDAP has a web application (human-friendly)

for every underlying web service (script-friendly).

You can use either one.

ERDDAP Doesn't Need a New Plug-in or Library

Plug-ins and libraries (client-side solutions) can be great.

They work at the point-of-need.

Sometimes they are the best or only solution.

However, they have disadvantages.

- Some users are reluctant to install them.
- Some users don't have permission to install any software.
- Some users will have difficulty installing them.
- Plug-ins may break when the application is updated.
- If there is a new version of the plug-in (e.g., with a bug fix),
it is impossible to get all users to update.

You probably don't even know who they are.

ERDDAP is a server-side solution.

It doesn't have these problems.

**One administrator updates ERDDAP and
the changes apply to all users, immediately.**

Problem #4: Incompatible Time Formats

Different datasets store date+time data in different formats, making them hard to compare.

("Jan 2, 1985", "02-JAN-1985", "1/2/85", "2/1/85", "1985-01-02", Year+DayOfYear+TimeOfDay, "seconds since 1970-01-01T00:00:00Z", "months since 1-1-1", ...)

ERDDAP's Solution:

- As a string, time is always in ISO 8601:2004(E) format
e.g., 2011-08-03T20:00:00Z
- As a number, time is always Unix Time
"seconds since 1970-01-01T00:00:00Z"
- The time zone is always Zulu (UTC, GMT)
- ERDDAP can convert string times to/from numeric times:
<http://coastwatch.pfeg.noaa.gov/erddap/convert/time.html>
(It's a web application and a web service!)

Problem #5: Lack of Metadata (information about the data)

- What is this dataset?
- Where is it from?
- What are the units?
- How can I find out more?

ERDDAP's Solution: More and Better Metadata

- Lets administrators add and modify the metadata
- Requires important metadata: title, summary, infoUrl, history
- Displays the metadata in many places (e.g., below forms, (?) pop-ups)
- Supports metadata standards:
 - Climate and Forecast (CF)
 - Unidata's Attribute Conventions for Dataset Discovery
 - Automatically generates (or uses custom) FGDC and ISO 19115
- Offers a service to convert UDUNITS to/from UCUM:
<http://coastwatch.pfeg.noaa.gov/erddap/convert/units.html>
(It's a web application and a web service!)

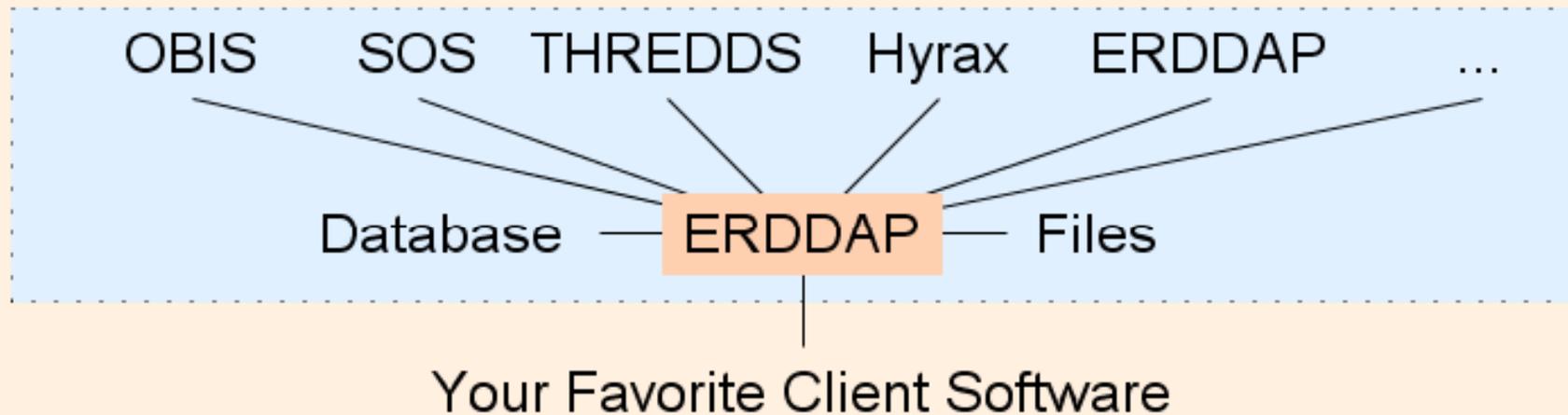
Summary of ERDDAP Features

- **Options to Search for Datasets of Interest**
 - Full-text search
 - Category-based (faceted) browsing
- **Standardized Request Format**
 - Web applications - web pages with forms for humans
 - Web services - one URL specifies the entire request
- **Standardized Response Formats**
 - Data file formats for your favorite application
 - Image file formats for custom graphs and maps
- **Standardized Time Formats**
 - Strings - ISO 8601:2004(E), e.g., 2011-08-03T20:00:00Z
 - Numbers - seconds since 1970-01-01T00:00:00Z
 - Time conversion application and service
- **Better Metadata**
 - So you can understand the data
 - Encourage standards - CF, ACDD
 - Generates FGDC and ISO 19115

ERDDAP

Environmental Research Division's Data Access Program

ERDDAP is a data server designed to help you get data from any data server into your favorite client software.



Try it: coastwatch.pfel.noaa.gov/erddap
Bob Simons <bob.simons@noaa.gov>
NOAA NMFS SWFSC ERD