The Unified Access Framework (UAF)

(a pretentious title for some simple ideas)

Kenneth.Casey@noaa.gov
Julie.Bosch@noaa.gov
Tina.Chang@noaa.gov
Scott.Cross@noaa.gov
Roy.Mendelssohn@noaa.gov
Steven.C.Hankin@noaa.gov
Jordan.Alpert@noaa.gov
Jeff.deLaBeaujardiere@noaa.gov
Ted.Habermann@noaa.gov
John.Relph@noaa.gov
Bob.Simons@noaa.gov
David.Neufeld@noaa.gov
Upendra.Dadi@noaa.gov
rsignell@usgs.gov
Phil.Cogbill@noaa.gov
Glenn.Rutledge@noaa.gov
Mike.Grogan@noaa.gov
Jeff.Budai@noaa.gov
Lewis McCulloch
Lewis.Mcculloch@noaa.gov
Matthew.Austin@noaa.gov

Presenter: Steve Hankin (PMEL),
for the UAF team

Enterprise-wide integration of data is a very hard problem!

(i.e. adopting shared IT strategies
to achieve data interoperability)

Why?

Within each field folks have developed IT solutions that make sense <u>to them.</u>

➔Getting people (and organizations) to change habits is difficult!

The NOAA ‘enterprise’

- weather forecast (time critical)
- fisheries management (regulatory concerns)
- nautical charting
- climate, ocean, atmosphere research

… the list goes on …

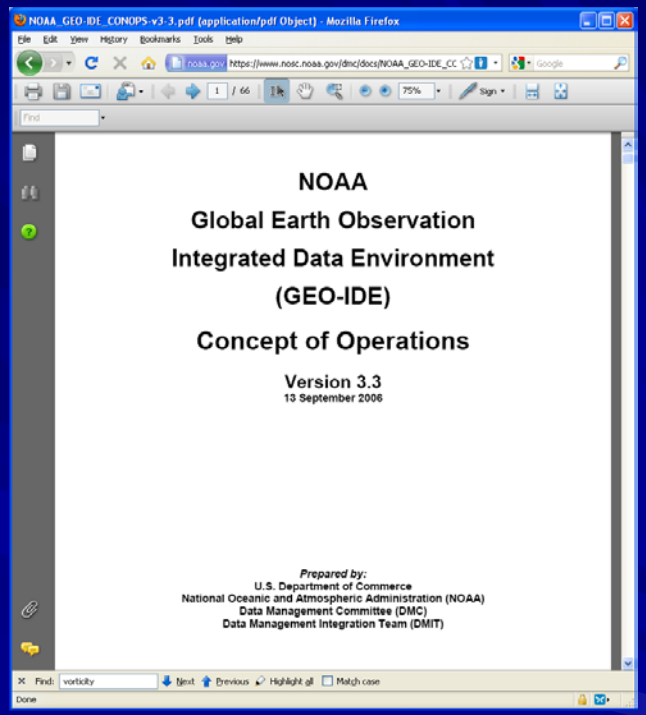
The accepted approach: build a “system of systems”

“Wrap” existing systems with loosely coupled, standardized services

➔ a Service Oriented Architecture

Global Earth Observation Integrated Data Environment

- NOAA's GEO-IDE Con-ops is such a plan
- 2005-06, 66 pages (by "DMIT")



Not the first nor the last such plan …

Seed funding finally available last year

Question:

How do you approach a very big problem with a very small amount of money?

Tried and true approach …

1. Assemble a writing team
2. Generate use cases
3. Define requirements
4. Write a Concept of Operations
5. … and an Implementation Plan
6. Assemble volunteer implementation teams

NOT !

An alternative (‘agile’) approach

Don't Solve Problems
-- Copy Success

"*Switch: How to Change Things When Change Is Hard*",
Chip and Dan Heath (psychologists), 2010

Projects: (too many to name)

Data formats: netCDF | GRIB | HDF | …

Service stack: netCDF-CF-DAP-THREDDS-WMS

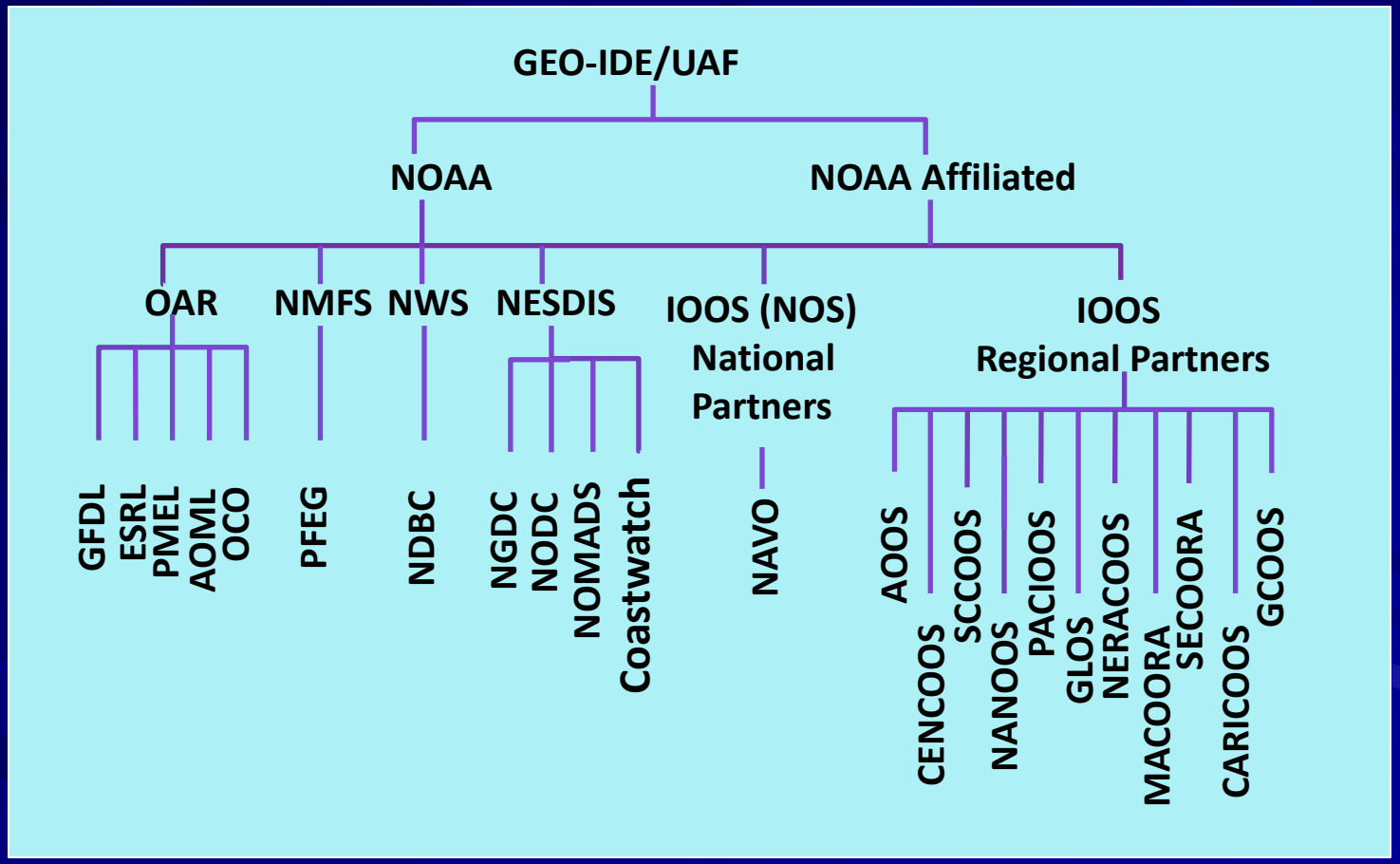
Applications: Matlab | ArcGIS | Ferret | NCL | NCO | GODIVA | W&Clim Toolkit | GrADS | Google Earth | IDV | LAS | ERDDAP | CDAT…

Users: (too many to name)

Rough spots in the real world

- a lot of “imperfect” (to be diplomatic) CF datasets
- too often unaggregated files rather than logical datasets
- metadata are often weak
- non-CF files (“trash”) mixed into the tree

UAF's foundation: the 'clean catalog'
(an XML tree of distributed, well-formed CF datasets)



Make clean virtual files using NcML tools

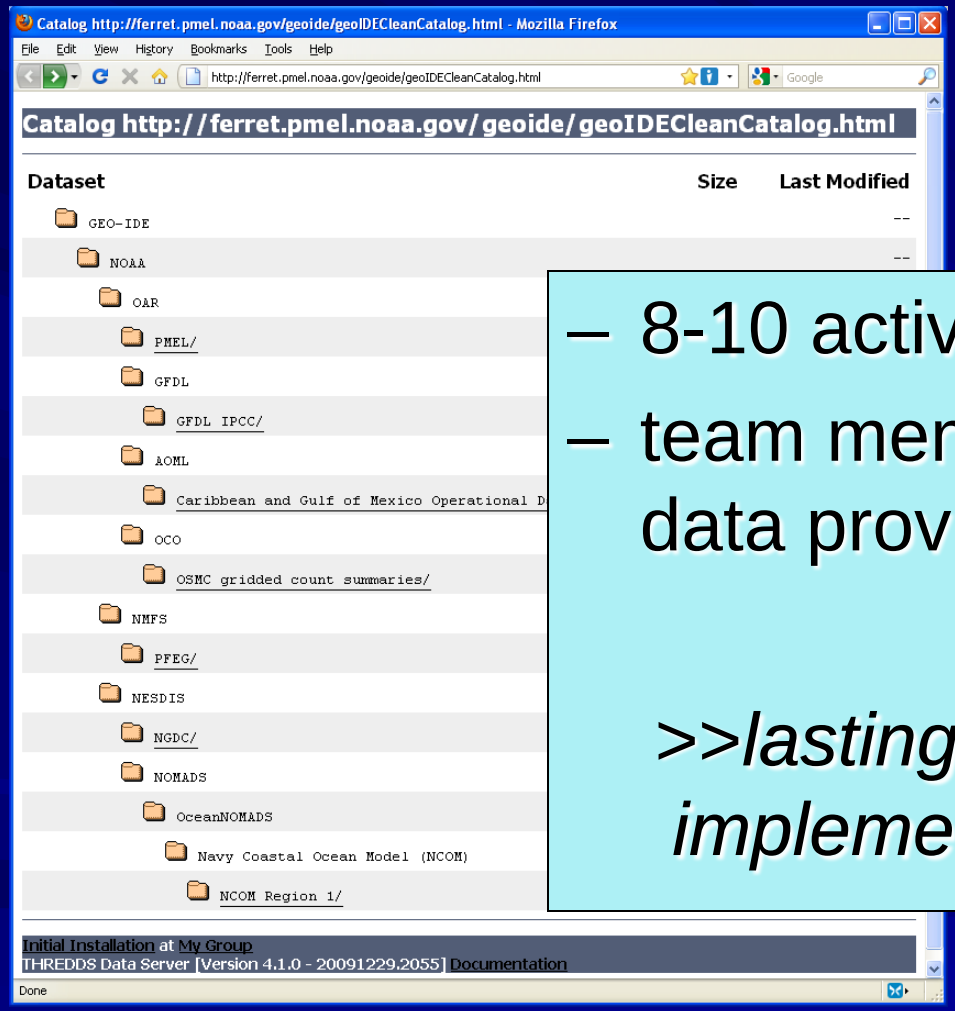
e.g. adding “standard_name” attribute to GRIB data

```
<variable name="vorticity">
  <attribute
    name="standard_name"
    value="atmosphere_absolute_vorticity“/>
</variable>
```

e.g. aggregation

```
<aggregation type="joinExisting“
            dimName="TimeAxis">
  <netcdf location=“year1.nc" ncoords=“365"/>
  <netcdf location=“year2.nc" ncoords=“365"/>
  <netcdf location=“year3.nc" ncoords=“365"/>
</aggregation>
```

The UAF Team



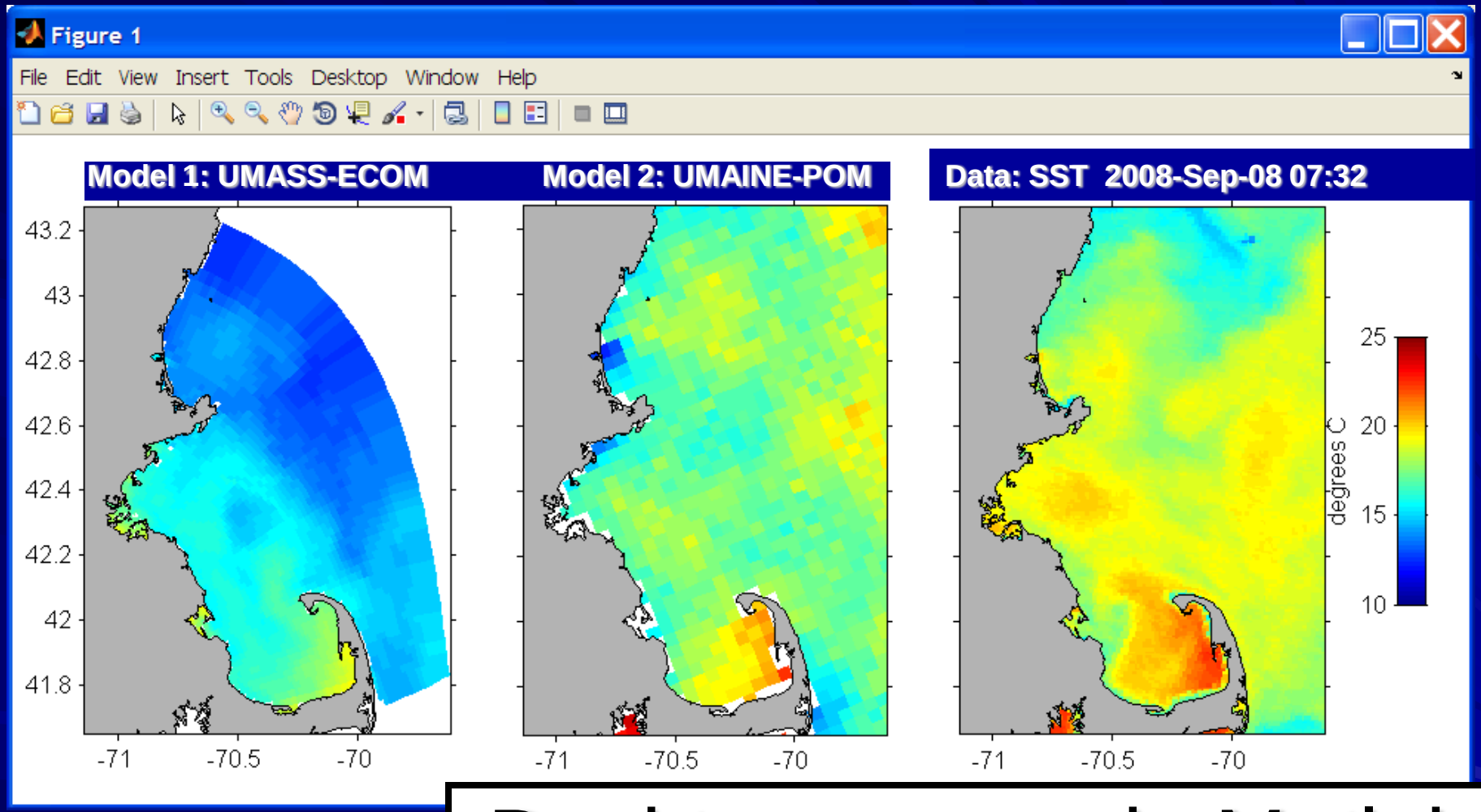
- 8-10 active “volunteers”
- team members as coaches to data providers

>>lasting solutions must be implemented by providers<<

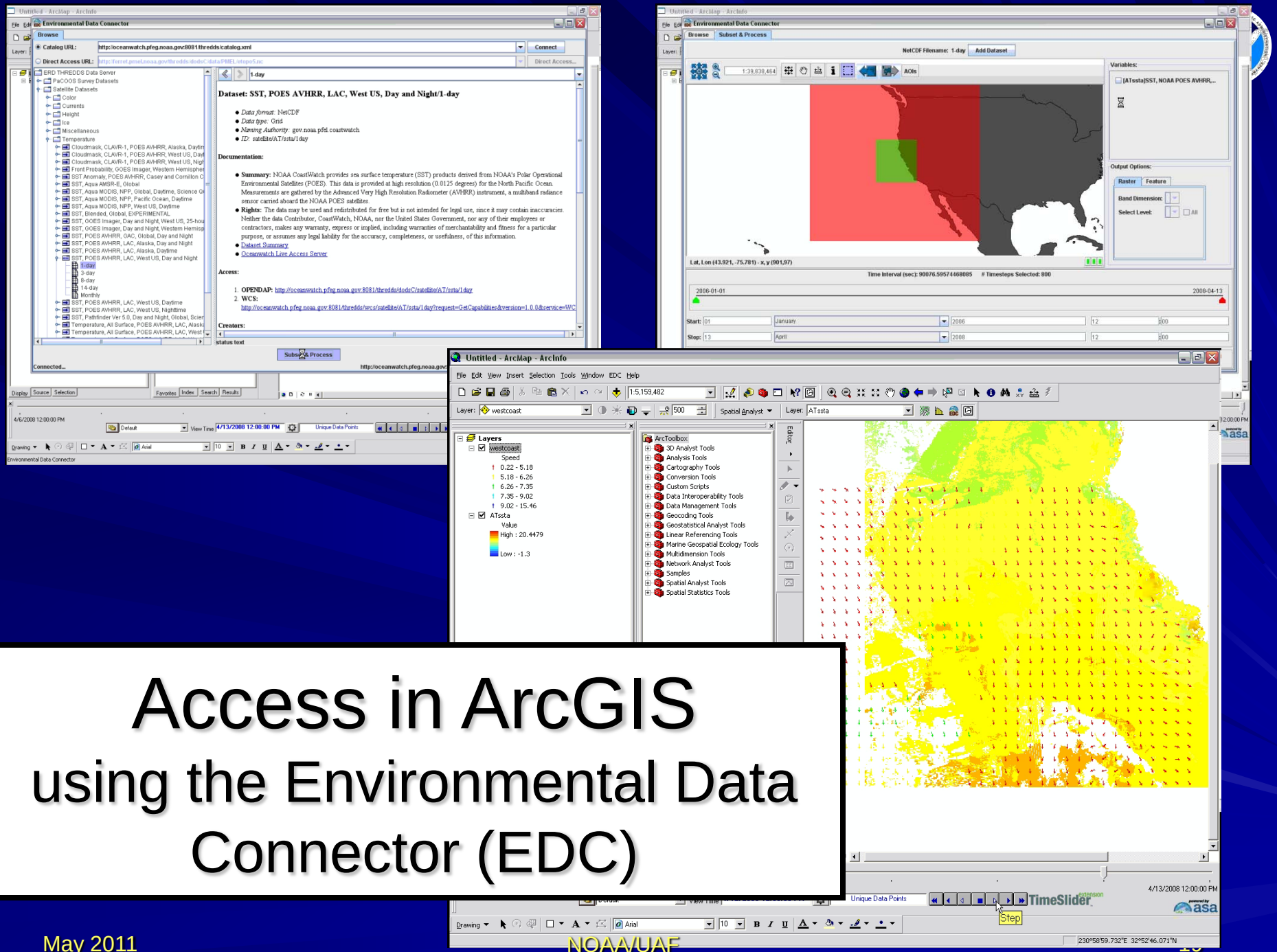
How to reach users?

Through their preferred tools

(without downloading files)



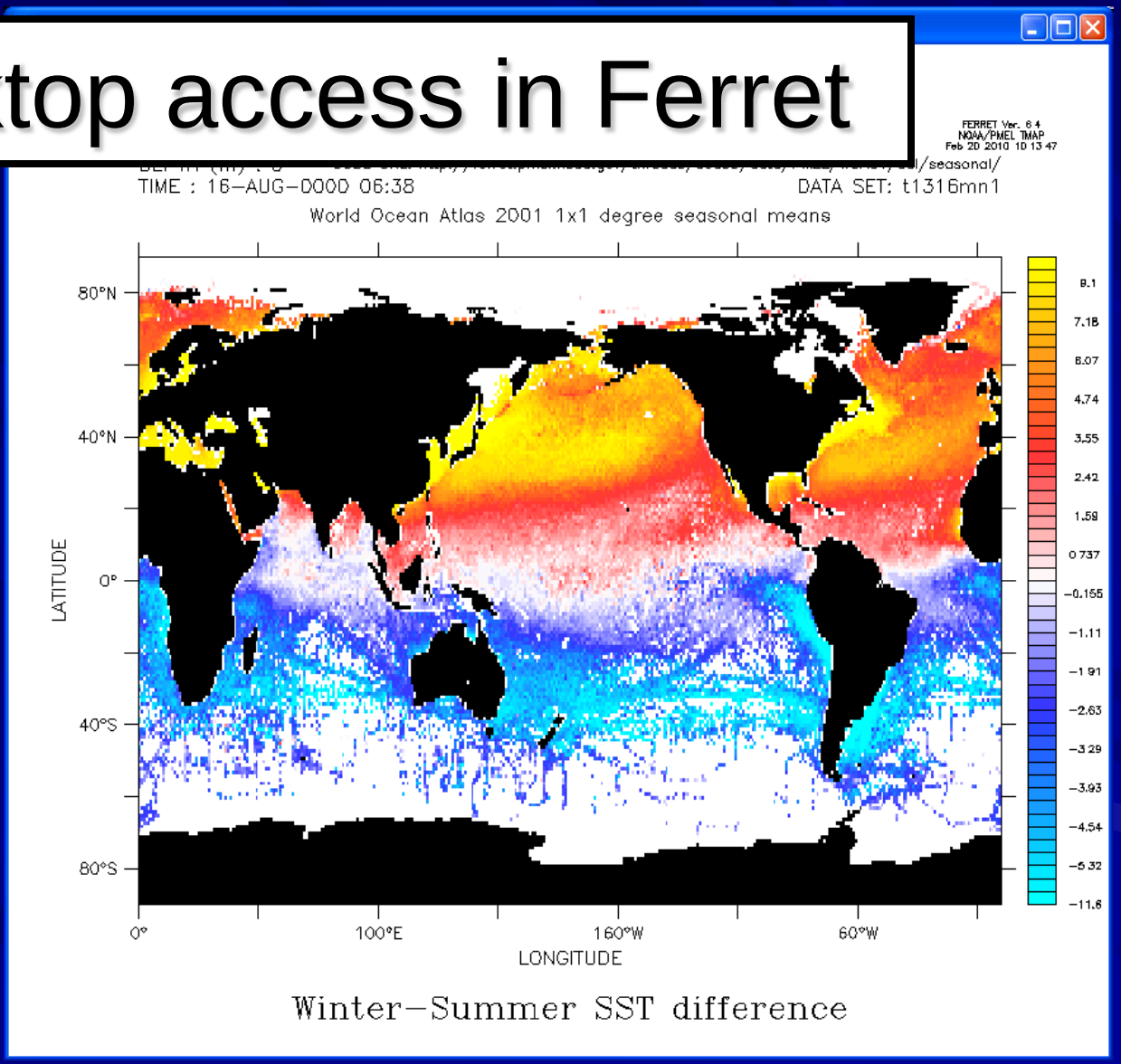
Desktop access in Matlab



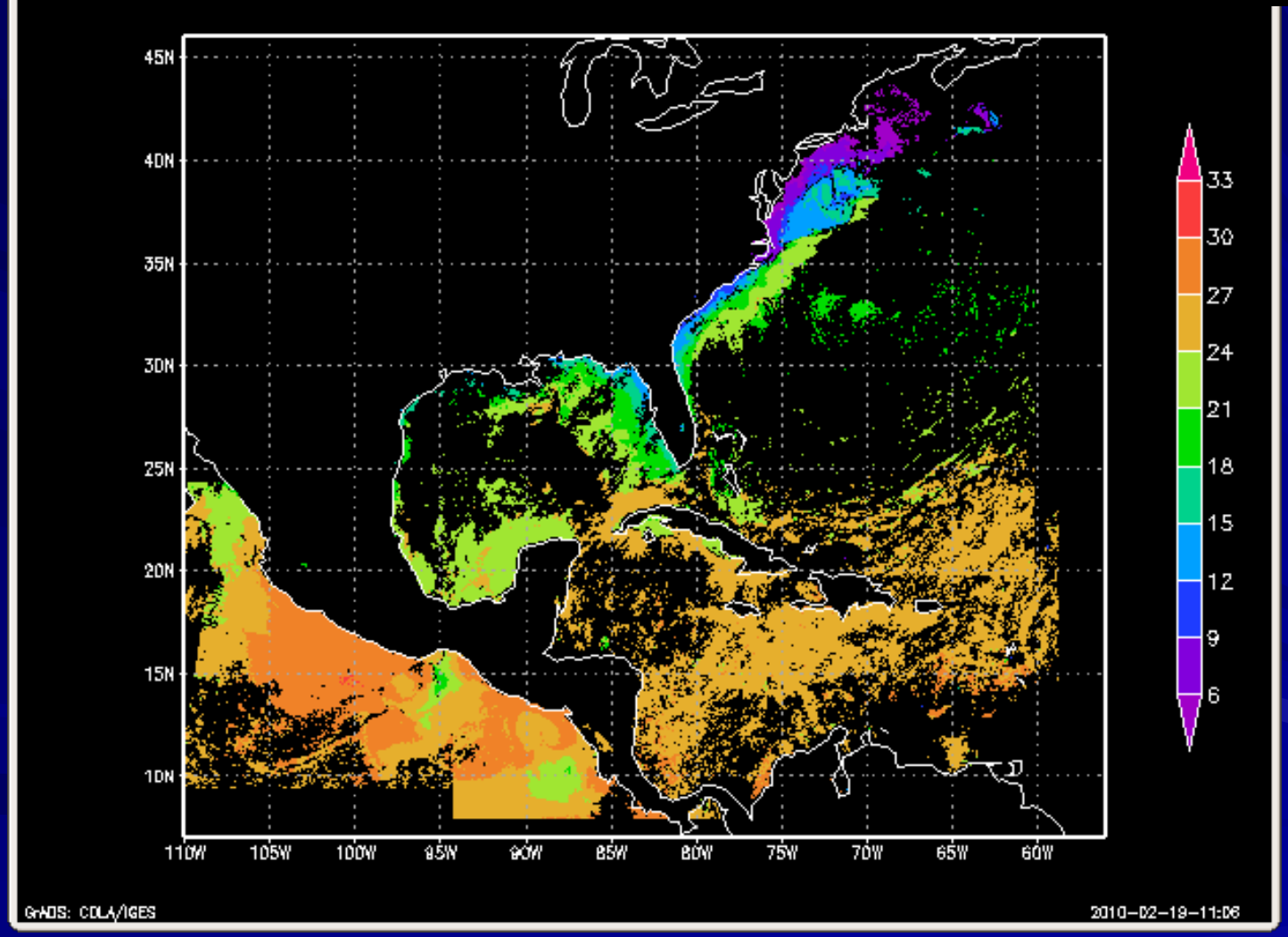
Access in ArcGIS

using the Environmental Data Connector (EDC)

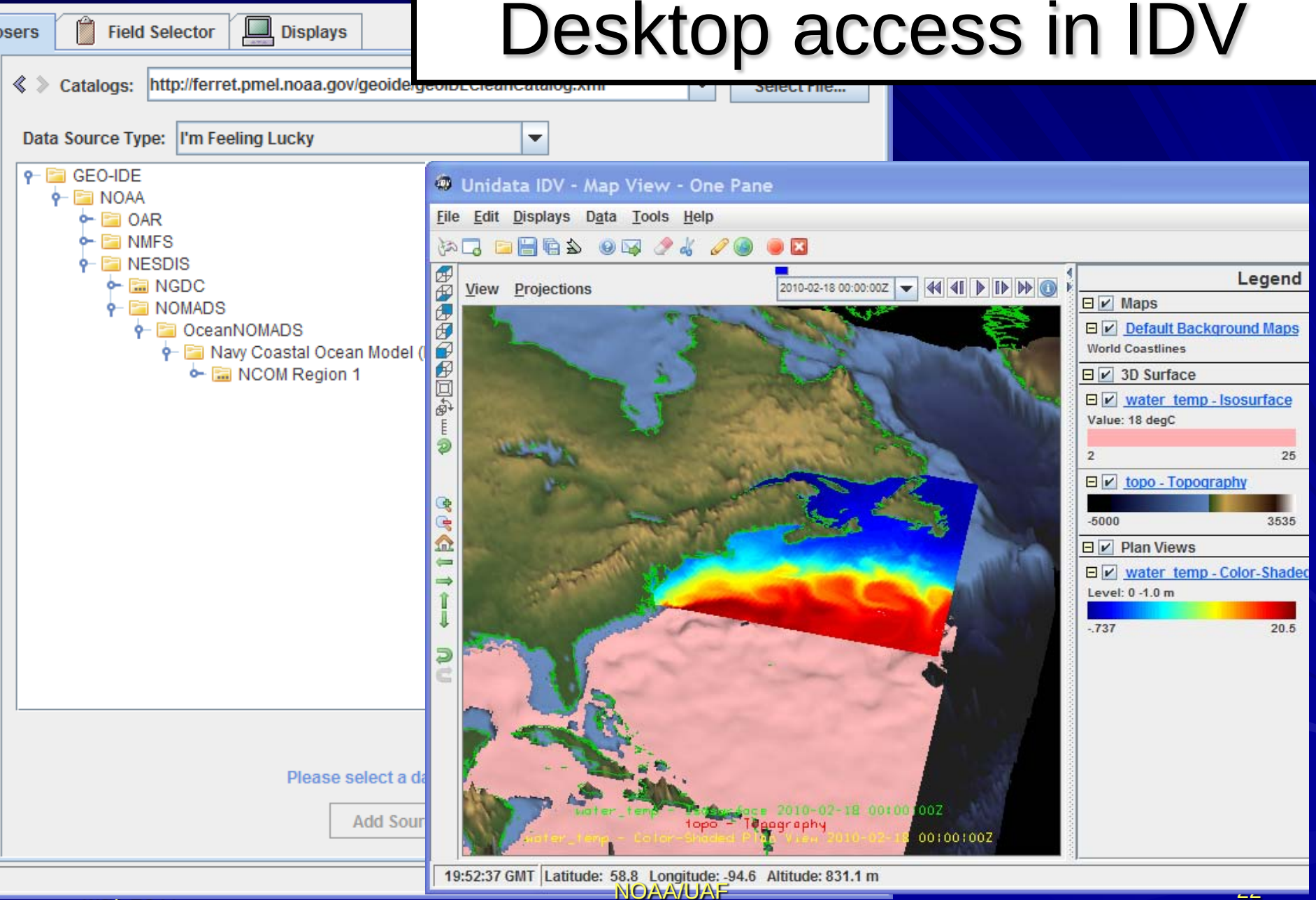
Desktop access in Ferret



Desktop access in GrADS



Desktop access in IDV



LAS OUTPUT - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://ferret.pmel.noaa.gov/geoideLAS/ProductServer.do?xml=<%3Fxml+version%3D"1.0"%3F><lasRequest

UAF/GEOIDE LAS | LAS OUTPUT

Ferret:

```
set data "http://ferret.pmel.noaa.gov/geoide/dodsC/las/id-ad3e40aa60
/data_ferret.pmel.noaa.gov_thredds_dodsC_data_PMEL_WOA01_ocl_monthly_t0112mn1.nc.jnl"

SET REGION/x="195.47":"286.88"/y="4.218":"54.843"/z="0":"0"/t="15-Jan-0001":"15-Jan-0001"
```

GrADS:

```
sdfopen http://ferret.pmel.noaa.gov/geoide/dodsC/las/id-ad3e40aa60
/data_ferret.pmel.noaa.gov_thredds_dodsC_data_PMEL_WOA01_ocl_monthly_t0112mn1.nc.jnl
set t 14 14
set level 1 1
set lat 4.218 54.843
set lon 195.47 286.88
```

Scripts

Matlab via netCDF-Java Toolbox:

```
url='http://ferret.pmel.noaa.gov/geoide/dodsC/las/id-ad3e40aa60
/data_ferret.pmel.noaa.gov_thredds_dodsC_data_PMEL_WOA01_ocl_monthly_t0112mn1.nc.jnl';
[t0112mn1,grid]=nj_subsetGrid(url,'t0112mn1',[195.47 286.88 4.218 54.843 ],'15-Jan-0001')
```

Matlab via loaddods:

```
% Region covered by this URL t=14-JAN:14-JAN z=0:0 y=4.5:54.5 x=195.5:286.5
loaddods(http://ferret.pmel.noaa.gov/geoide/dodsC/las/id-ad3e40aa60
/data_ferret.pmel.noaa.gov_thredds_dodsC_data_PMEL_WOA01_ocl_monthly_t0112mn1.nc.jnl?t0112mn1[14,14][1,1][95,145]
[196,287])
```

Done

Date :

Minimum

Maximum

Start date: 0001 Jul

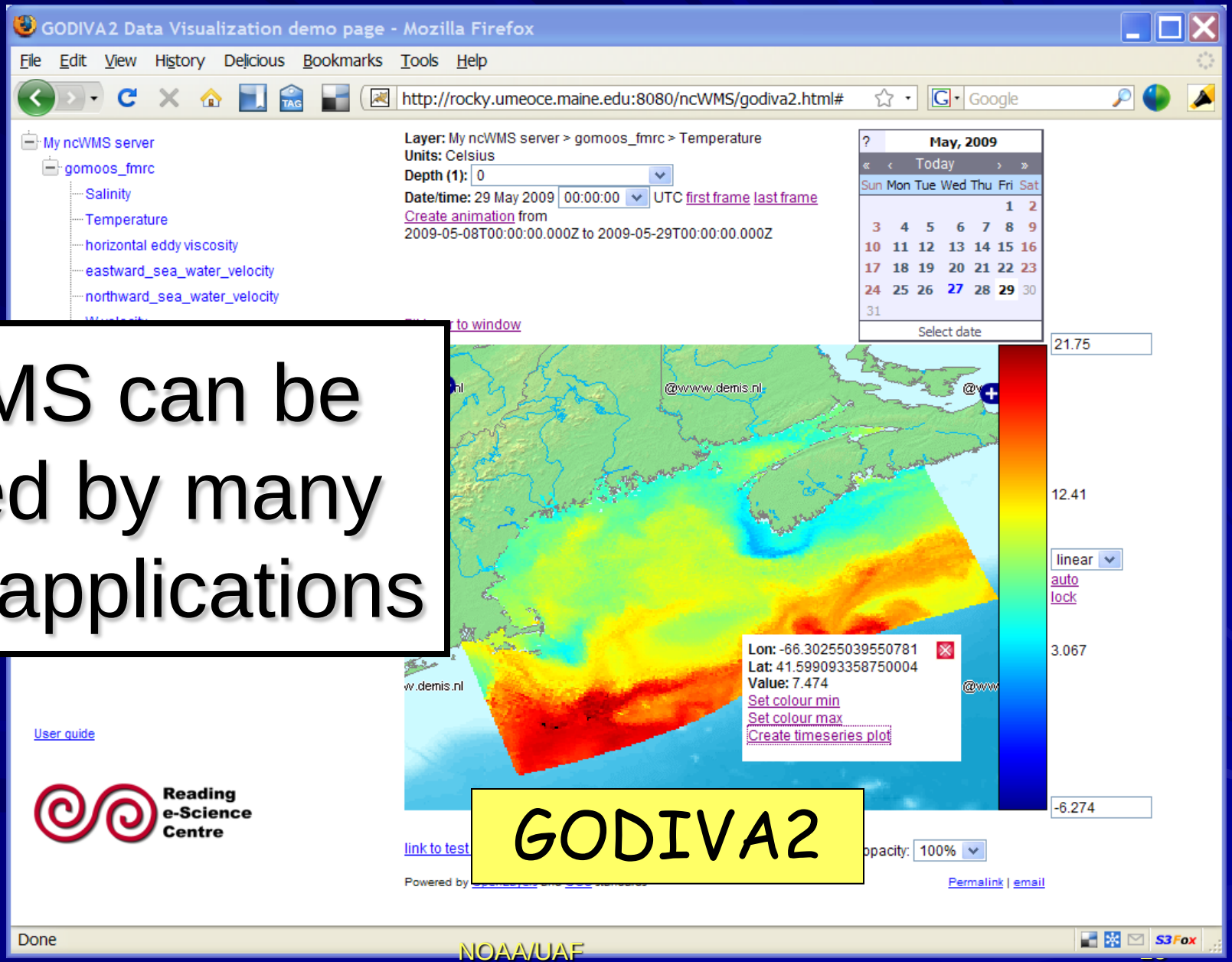
Start date: 0001 Oct

Done | Windows | Inbox for steven.c.h... | LAS vizGal - Mozilla Fi... | Compose: Re: UAF LAS | Hankin-AGU-20...

Live Access Server (LAS)

NOAA/UAF

WMS can be used by many GIS applications



GODIVA2

ERDDAP Easier access to scientific data

ERDDAP > griddap > Data Access Form

Dataset Title: **SST, Blended, Global, EXPERIMENTAL (Monthly Composite)**
Dataset ID: erdBAsstamday RSS
Institution: NOAA CoastWatch, West Coast Node
Information: Summary | Variables | Background | Make A Graph

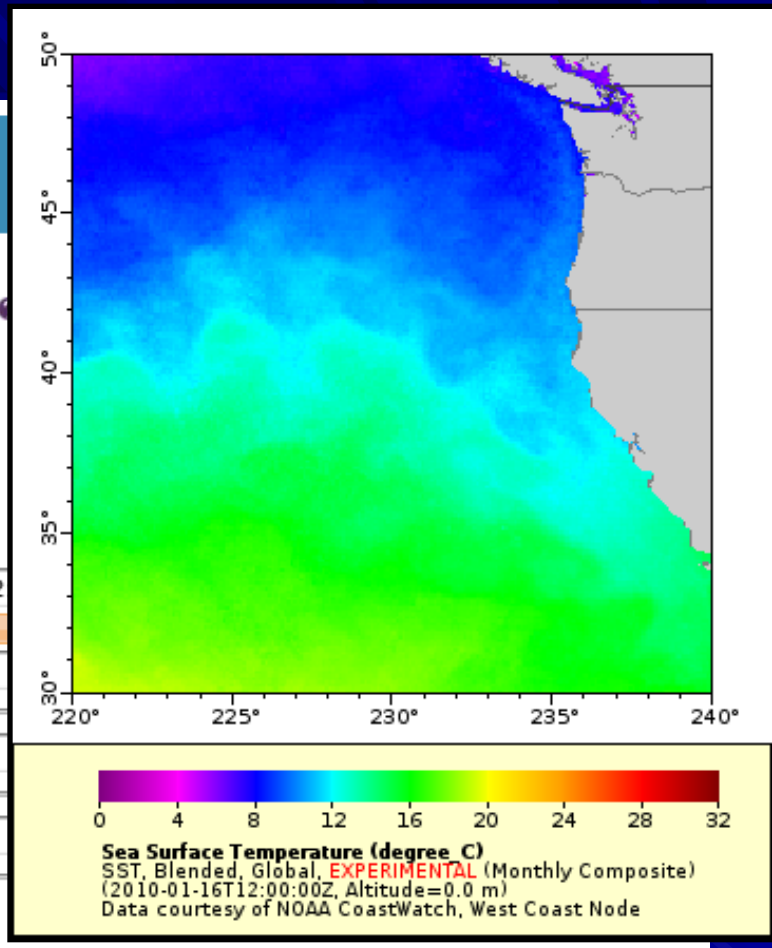
Dimensions	Start	Stride	Stop
time (Centered Time, UTC)	2010-01-16T12:00:00Z	1	2010-01-16T12
altitude (m)	0.0	1	0.0
latitude (degrees_north)	30	1	50.0
longitude (degrees_east)	220	1	240.0

Grid Variables (which always also download all of the dimension variables)
sst (Sea Surface Temperature, degree_C)

File type: .png - Download a standard, medium-sized .png image file with a map of the data you selected.

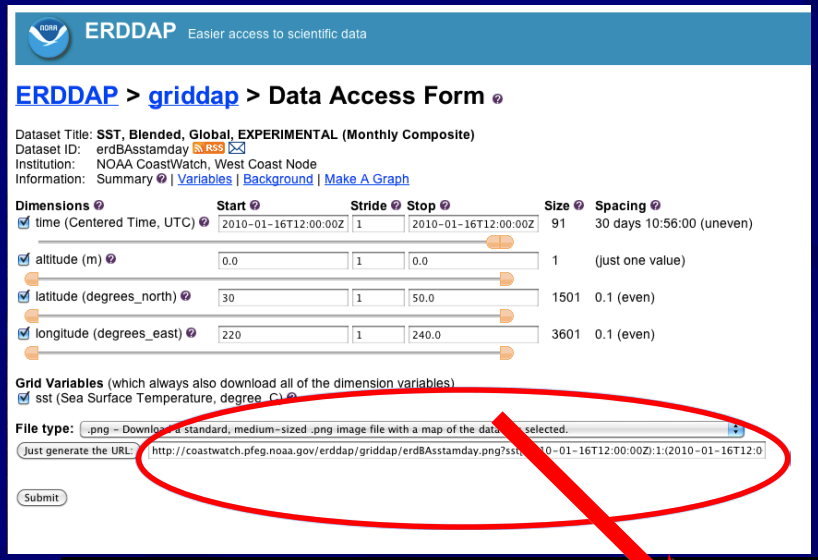
Just generate the URL: http://coastwatch.pfeg.noaa.gov/erddap/griddap/erdBAsstamday.png?sst[(2010-01-16T12:00:00Z):1:(2010-01-16T12:0

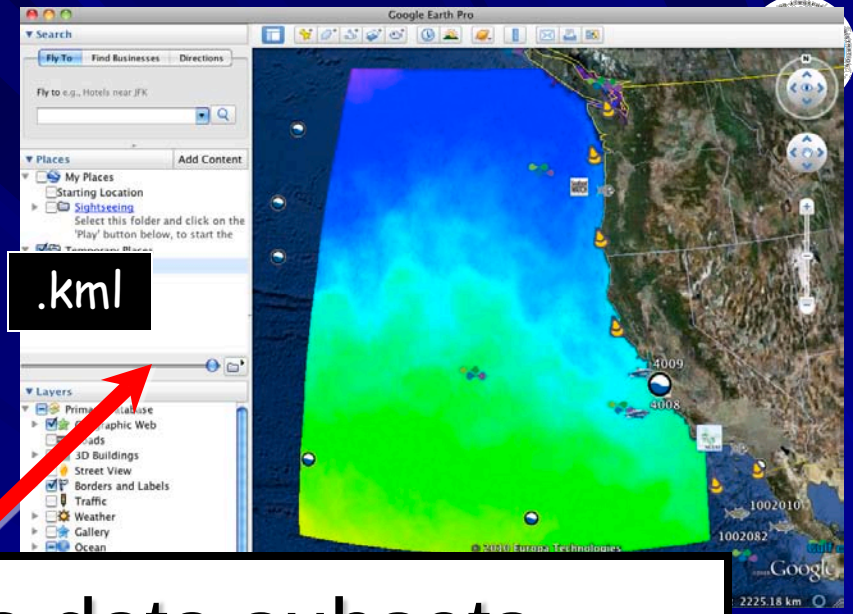
Submit



Browser access using ERDDAP

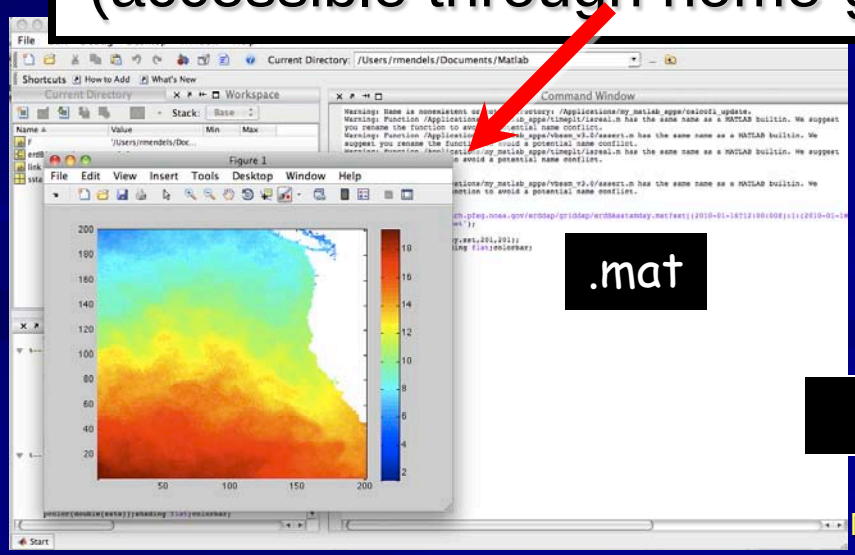
…and from there …

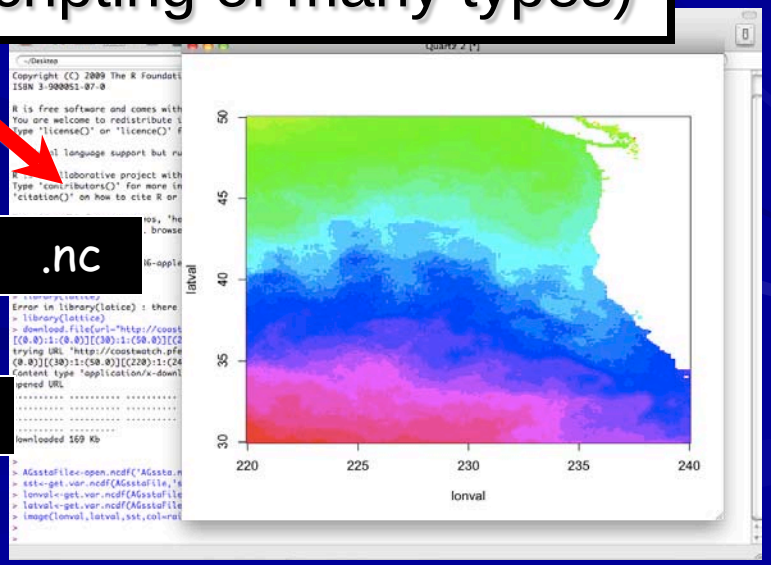




REST URL access to data subsets in several popular formats

(accessible through home-grown scripting of many types)





also .csv

NOAA/UAF

the list goes on …

- Weather and Climate Toolkit
- NCL (NCAR)
- NCO
- CDAT / UVCDAT (DOE)
- ncview
- Giovani (NASA)
- generic WCS & WMS applications
- …

Global Earth Observation - Integrated Data Environment | Access Gridded Data - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://geo-ide.noaa.gov/access.html

CVS PMEL LAS Projects pending Webster Ferret Google NOAA Directory trip

Global Earth Observation Integrated Data Environment

Unified Access Framework for Environmental Data

Home | Access Data | Find Data | Contribute Data | Learn About UAF | UAF in Action | GEO-IDE Wiki | Contact Us

- Home
- Access Data
- Find Data
- Contribute Data
- Learn About UAF
- UAF in Action
- GEO-IDE Wiki
- Contact Us

Interoperability Provided By:

Access Data

One of the primary benefits of the Unified Access Framework is the ease and flexibility of accessing a large amount of data in ways that YOU, the user, wants. Many options are already available to take advantage of the Unified Access Framework. Click on one of the logos below to learn more...

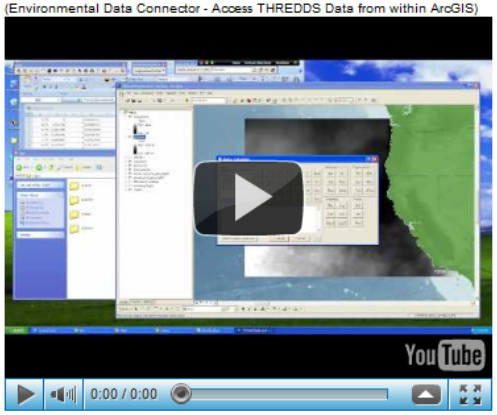
THREDDS | LAS | ERDDAP | IDV | NCO | ncBrowse

Ferret | ArcGIS | MATLAB | GrADS | Unidata Tools

ArcGIS

- Learn about ArcGIS

(Environmental Data Connector - Access THREDDS Data from within ArcGIS)

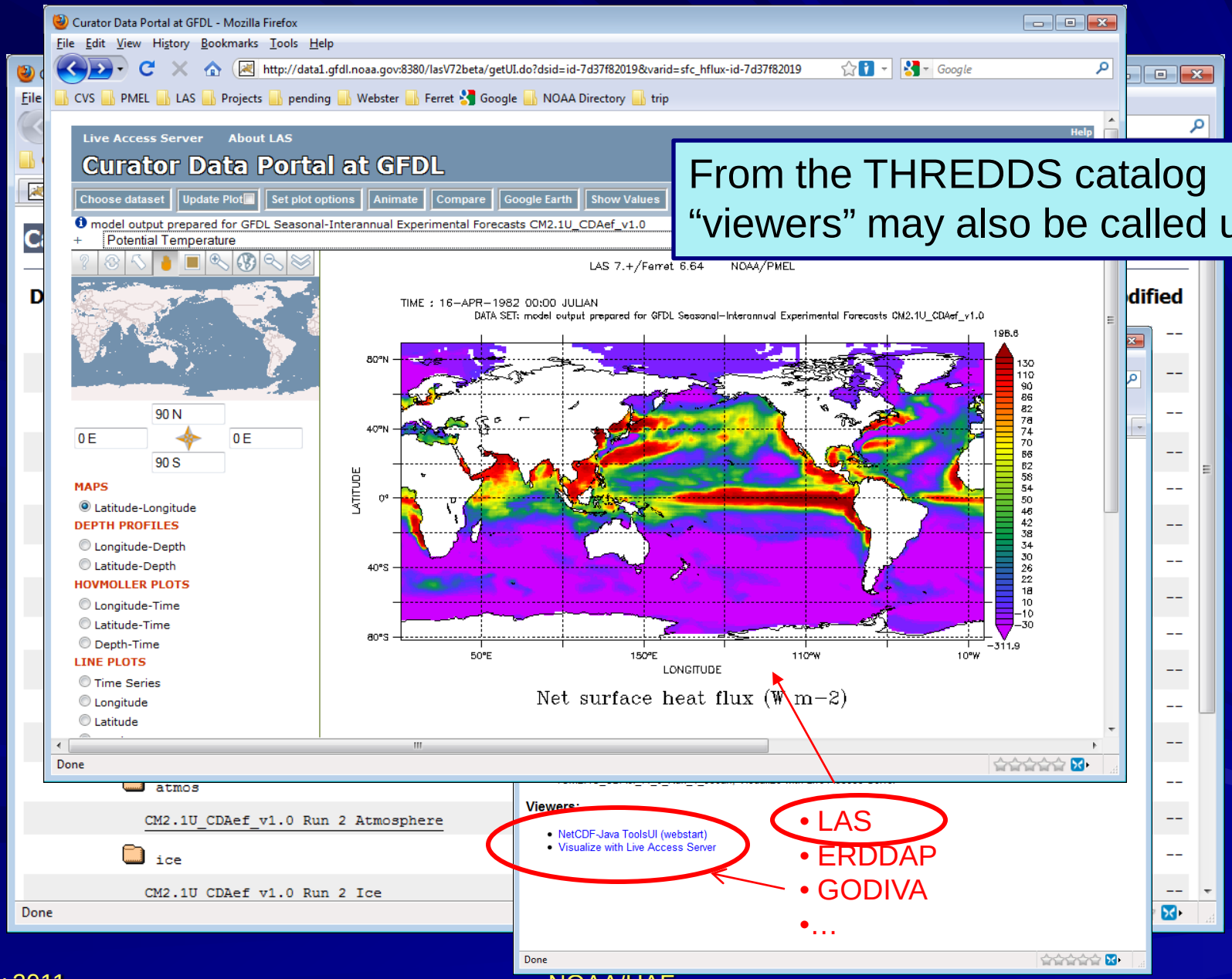


(Sample Screenshot)

Privacy Policy | Disclaimer | Site Map | Contact Webmaster

Read i1.ytimg.com

From the THREDDS catalog "viewers" may also be called up



Metadata inside files has become accessible from THREDDS catalogs

Catalog Services - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ngdc.noaa.gov/thredds/sstCatalog.html?dataset=SST-Aerosol-.

CVS PMEL LAS Projects pending Webster Ferret Google NOAA Directory trip

Curator Data Portal at GFDL | Catalog Services

National Geophysical Data Center

THREDDS Data Server

Catalog http://www.ngdc.noaa.gov/t

Dataset: SST Catalog/SST Aerosol A

- *Data type:* GRID
- *ID:* SST-Aerosol-Agg

Access:

1. **OPENDAP:** /thredds/dodsC/sst-aerosol-aggregation
2. **HTTP Server:** /thredds/fileServer/sst-aerosol-aggregation
3. **NCML:** /thredds/ncml/sst-aerosol-aggregation
4. **UDDC:** /thredds/uddc/sst-aerosol-aggregation
5. **ISO:** /thredds/iso/sst-aerosol-aggregation

Viewers:

- Integrated Data Viewer (IDV) (webstart)
- NetCDF-Java ToolsUI (webstart)

Done

Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ngdc.noaa.gov/thredds/ncml/s

CVS PMEL LAS Projects pending Webster Ferret Google NOAA Directory

Curator Data Portal at GFDL | http://www.ngdc....sol-aggregation

This XML file does not appear to have any style information associated with it. The document tree is shown below.

<netcdf location="file:sst-aerosol-aggregatio
<attribute name="Conventions" value="
<attribute name="standard_name_vocab
<attribute name="naming_authority" val
<attribute name="geospatial_lon_units"
<attribute name="geospatial_lat_units" v
<attribute name="geospatial_lon_resolut
<attribute name="geospatial_lat_resoluti
<attribute name="cdm_data_type" value
<attribute name="creator_url" value="h
<attribute name="time_coverage_duratio
<attribute name="history" value="2010-
NPR.ASWK.NP.D1026605 into netCDF
<attribute name="Metadata_Convention
<attribute name="metadata_url" value="
/fgdc.jsp?id=gov.noaa.class:AERO100"/>
<attribute name="id" value="gov.noaa.c
<attribute name="title" value="Aerosol
<attribute name="summary" value="The
optical thickness based on one week's wo
channel 1 optical thickness retrievals from
thickness, or optical depth, is a dimensionl
beam of radiation undergoes as it passes th
function of the density, composition, tempe
defined as suspensions of liquid droplets o
sand, volcanic ash, sea spray and smog."/>
<attribute name="keywords" value="cli
SCIENCE > Atmosphere > Aerosols > A
Atmosphere > Atmospheric Radiation > C
Atmosphere > Clouds > Cloud Optical De
Optics > Optical Depth, CIESIN > Enviro
Aerosols, atmosphere_optical_thickness_
<attribute name="keywords_vocabulary
Earth Science Keywords, CIESIN Indexin
Standard Name Table"/>
<attribute name="creator_name" value=

Done

Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ngdc.noaa.gov/thredds/iso/sst-ae

CVS PMEL LAS Projects pending Webster Ferret Google NOAA Directory trip

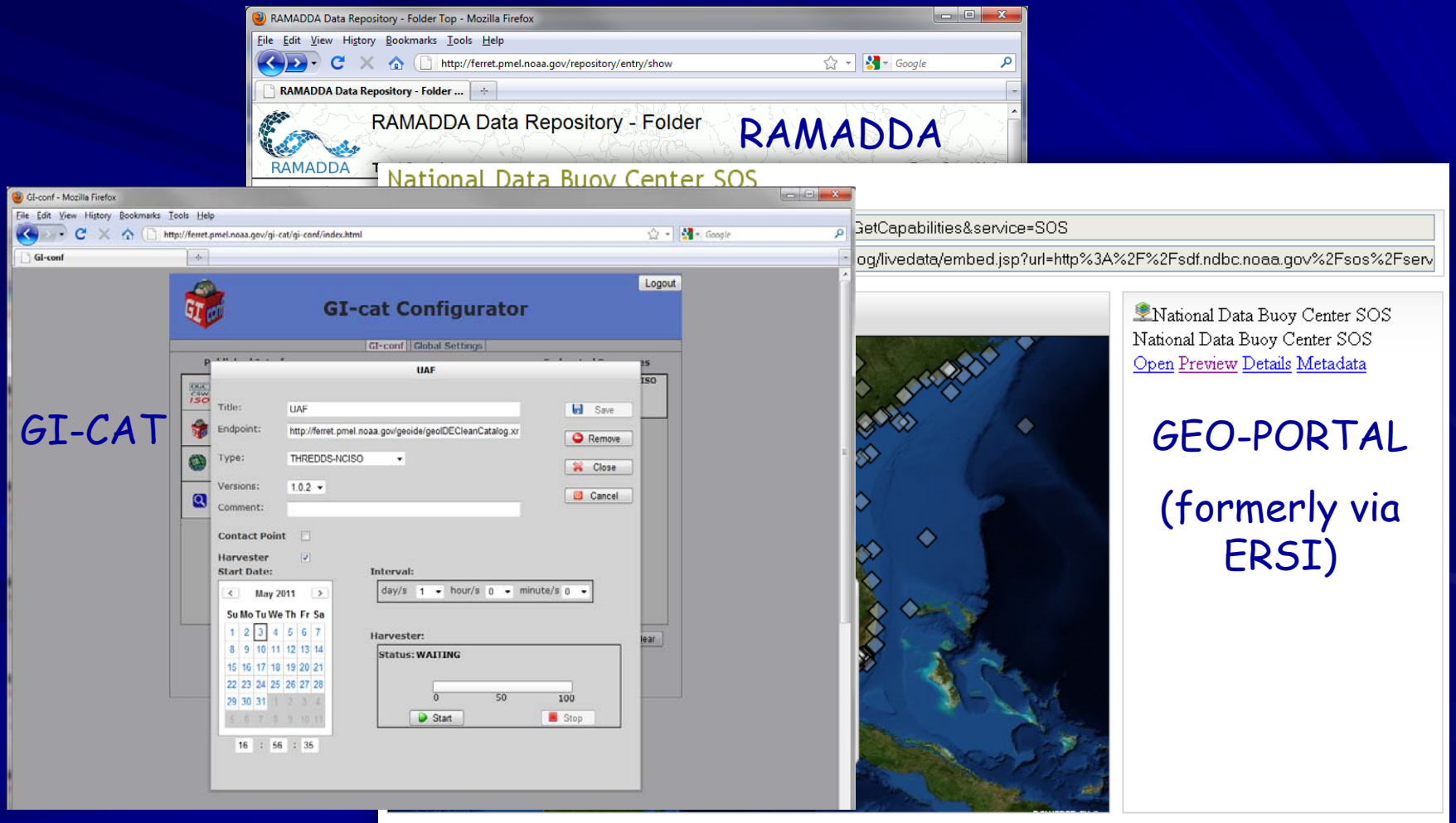
This XML file does not appear to have any style information associated with it. The document tree is shown below.

<gmi:MI_Metadata xsi:schemaLocation="http://www.isotc211.org/2005/gmi C:/DOCUME~1/haber
/MYDOCU~1/NOAA/Metadata/ISOSTA~1/ISO191~3/19139S~1/gmi/gmi.xsd">
<gmd:fileIdentifier>
<gco:CharacterString>gov.noaa.class:AERO100</gco:CharacterString>
</gmd:fileIdentifier>
<gmd:contact gco:nilReason="unknown"/>
<gmd:dateStamp gco:nilReason="unknown"/>
<gmd:spatialRepresentationInfo>
<gmd:MD_Georectified>
<gmd:numberOfDimensions gco:nilReason="unknown"/>
<gmd:axisDimensionProperties>
<gmd:MD_Dimension>
<gmd:dimensionName>
<gmd:MD_DimensionNameTypeCode codeList="http://www.isotc211.org
/2005/resources/codeList.xml#MD_DimensionNameTypeCode"
codeListValue="column">column</gmd:MD_DimensionNameTypeCode>
</gmd:dimensionName>
<gmd:dimensionSize gco:nilReason="unknown"/>
<gmd:resolution>
<gco:Measure>1</gco:Measure>
</gmd:resolution>
</gmd:MD_Dimension>
</gmd:axisDimensionProperties>
<gmd:axisDimensionProperties>
<gmd:MD_Dimension>
<gmd:dimensionName>
<gmd:MD_DimensionNameTypeCode codeList="http://www.isotc211.org
/2005/resources/codeList.xml#MD_DimensionNameTypeCode"
codeListValue="row">row</gmd:MD_DimensionNameTypeCode>
</gmd:dimensionName>
<gmd:dimensionSize gco:nilReason="unknown"/>
<gmd:resolution>
<gco:Measure>1</gco:Measure>
</gmd:resolution>
</gmd:MD_Dimension>
</gmd:axisDimensionProperties>

Done

Our current efforts …

1. evaluating ‘mature’ discovery portals



How to obtain discovery metadata?

Crawl the UAF clean catalog with

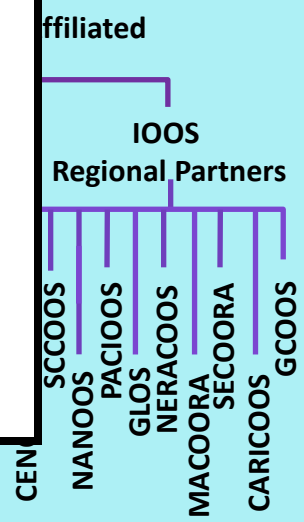
- RAMADDA
- GI-CAT
- ncISO/NGDC
  - also create Web Accessible Folders to join into other frameworks

Our current efforts …

2. enhancing the “catalog cleaner” (another THREDDS crawler)

The clean catalog:

- **U**niform services and viewers
- data access via OPeNDAP as needed
- auto-synchronization (cron-style)
- join-new and join-existing aggregations



Our current efforts …

3. *in situ* obs collections (“1d grids”)

(development work …)

1. CF Discrete Geometries spec

now official

2. ncStream (‘cdmRemote’) (Caron)
3. ERDDAP (tableDAP REST), LAS (anal & viz)
   and after these …
4. … IOSP for data base access
5. … NCML aggregation of 1d file collections

Summary

The UAF approach is

- simple, open, cheap
- compatible with ESGF, NOMADS, IOOS, Ingrid, Giovani, OGC (recently), …

Should it be broader than a NOAA project?

Many OPeNDAP servers provide little user documentation.

➔ Send us docs to host on your CF app

If UAF were at a non-agency domain, would you hang your open THREDDS catalog on the UAF tree?

Thank you