



# Emerging Tools for Distributed Data Access and Collaborations

**Glenn K. Rutledge**

National Oceanic and Atmospheric Administration

National Climatic Data Center

13<sup>th</sup> Federation Assembly Meeting

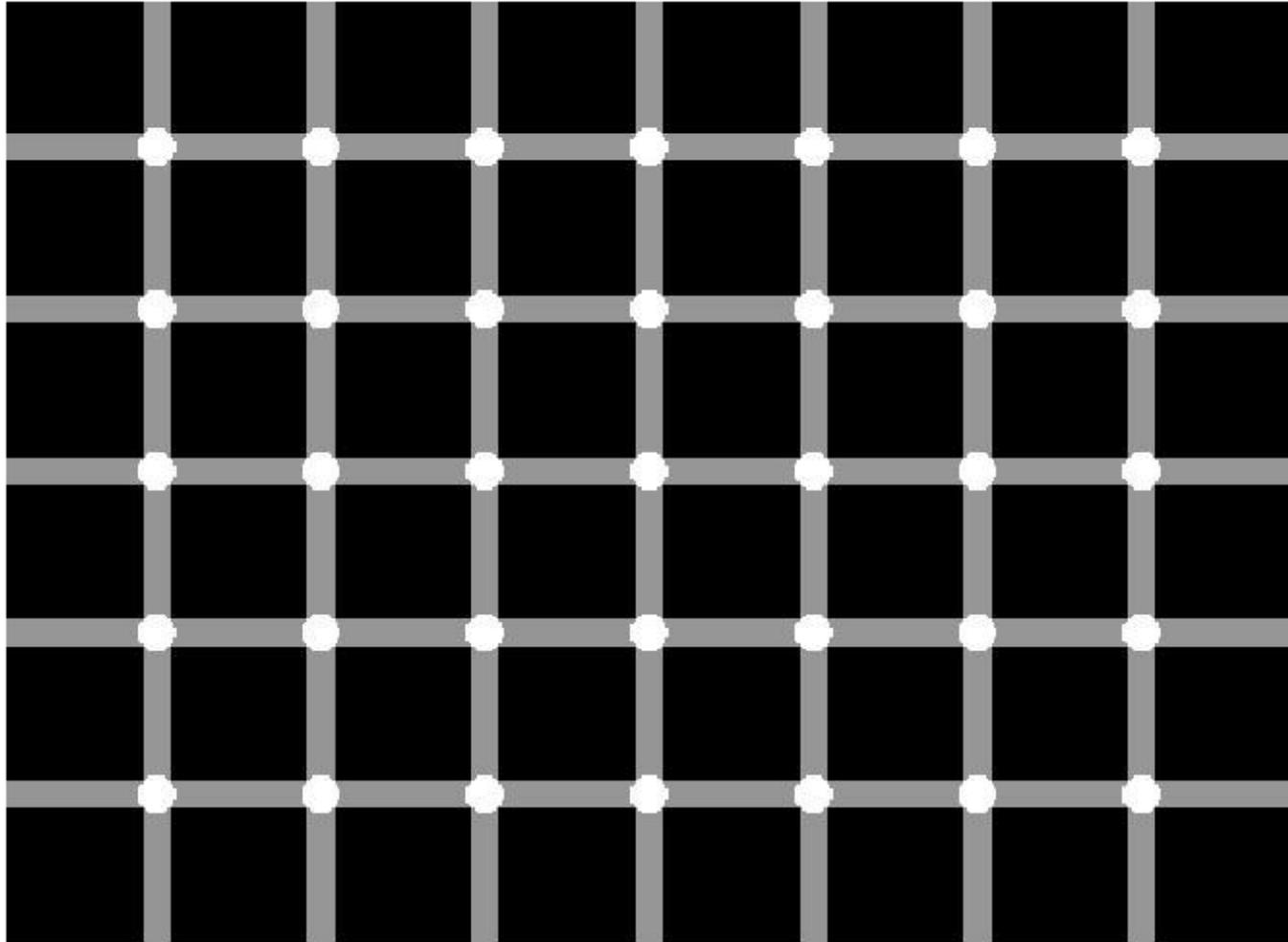
Earth Systems Information Partners

Asheville, NC August 18, 2004



- Earth System Interoperability
  - Science Exploration at the Data Level
- Metadata, Catalogs, & Ontologies
- Tools and Programs
  - Too many to count?
- My World: NOMADS
- Next Steps to Collaborations

Count the black dots....



XML  
SOAP  
OWL  
OPeNDAP

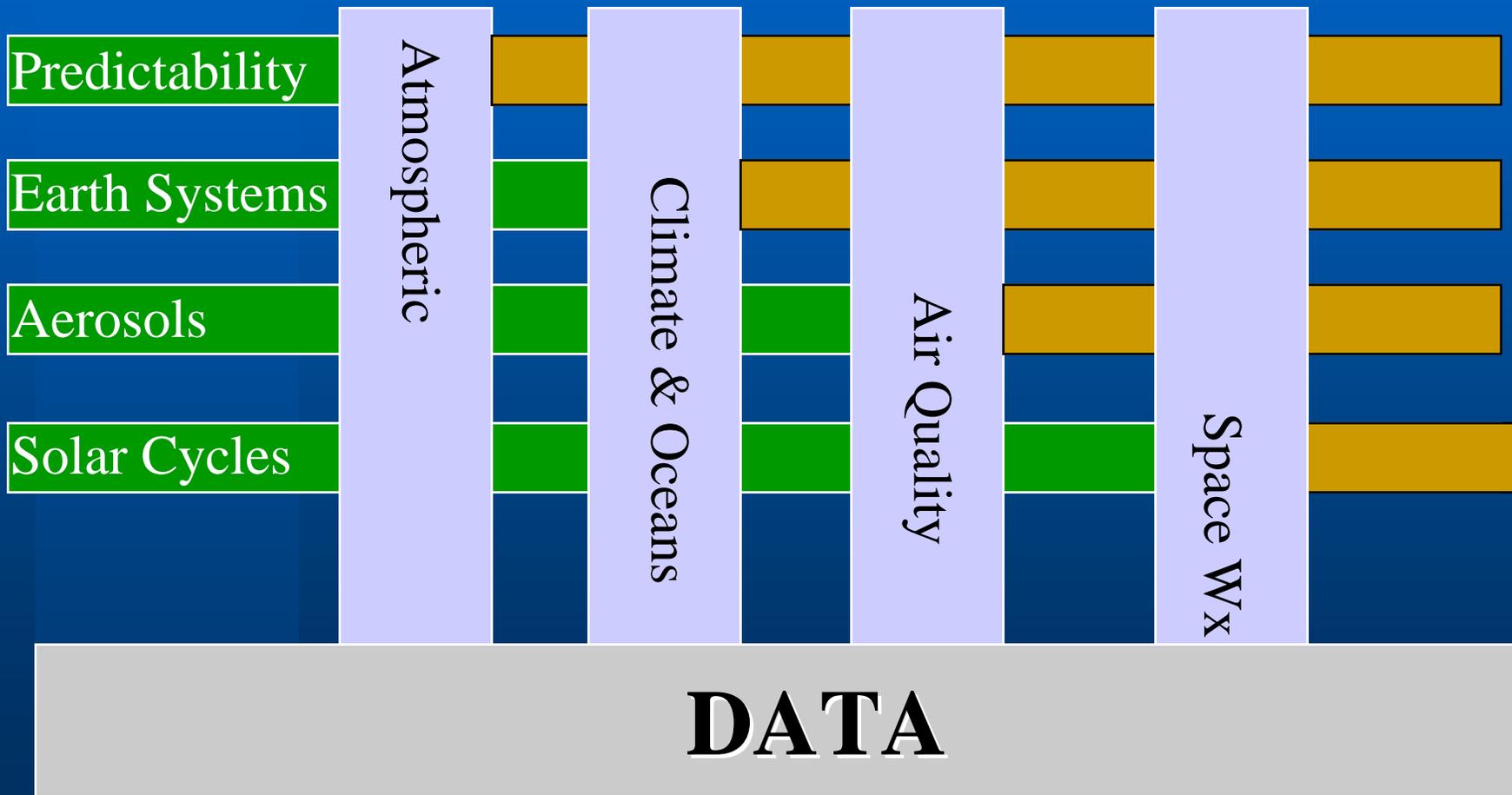
Data Interoperability...  
A moving target?  
Yes.

ESML  
http/TCP  
Globus  
SWEET

## *Science Exploration at the Data Level*

- What are the goals facing the GeoScience community?
  - Is it just **access** to high volume data (satellite, radar, and model)?
- How will Agencies and Institutions address interoperability?
  - Should it be system, data or both?
- Have the scientific requirements been adequately defined?
  - Do top down approaches adequately promote science?
- How can Agencies and institutions develop partnerships while allowing for attribution, with diverse goals and agendas?
- **Data** interoperability is the key: Scientific Data Stewardship

# *Program Management at the Data Level*



- Operational Forecasting-
  - Ensemble Predictions: flow-dependant prediction of weather and climate risk- nowcasting, medium range and seasonal.
- Atmospheric and Oceanic Research-
  - Scalar and Vector processing and Workstation models
  - Model output statistics; data assimilation techniques
- Global Climate Change and Advanced Analysis-
  - Clouds, initial conditions, true coupled simulations.
  - Long term climate monitoring: in-situ analysis, trends, data homogeneity, extremes, downscaling, reducing uncertainty...
  - On-demand Data Mining and Product Generation.

- *Data systems based on the integration of independently developed system elements offer many more opportunities than more traditional centrally developed ones.*

Peter Cornillon URI  
...et al.

## *Collaborations: How do we get there?*

- Data transport is being actively pursued: OPeNDAP, SOAP, ...
- Earth System Partners need to be able to find and use various data sets, wherever they may be, whatever format...
- THREDDS can provide dynamic access and generate catalogs
- GCMD is a major resource for metadata management for the entire GeoSciences community- this activity must evolve!
- Ontology projects such as SWEET in conjunction with THREDDS and GCMD can provide individual data sources, data variables and metadata management for the community.



## *Ontology: SWEET*

- The Semantic Web for Earth and Environmental Terminology (SWEET) project provides a common semantic framework for various Earth science initiatives.
- The semantic web is a transformation of the existing web that will enable software programs, applications, and agents to find meaning and understanding on web pages.
- SWEET developed these capabilities in the context of finding and using Earth science data and information.

- Pare down large file sizes of high resolution data and products.
- (re-) Group different data sets to create needed products – such as initialization files for model development, analysis, and intercomparison.
- Subset the data sets in parameter space
- Subset the data sets in physical space
- Subset the data sets in temporal space

## *Tools for Users (cont.)*

- Data extraction for the generation of products “on-demand”.
- Advanced data mining algorithms for pre-generation, or executed by (authorized) users also on-demand.
- Access to mined physical processes or signatures thru data mining.
- Search and location tools and metadata management.



# Metadata and Catalog Programs

**Just several programs addressing the data access, description, and search activities:**

- **CLASS** Comprehensive Large-Arary Stewardship System
- **DAAC** Distributed Active Archive Centers
- **DIMES** DIstributed MEtadata Server
- **DLESE** Digital Library for Earth System Education
- **ECHO** The EOS ClearingHouse (middleware)
- **ESIP** Earth Science Information Partners
- **FGDC** Federal Geographic Data Committee
- **FIND** Federation Interactive Network for Discovery
- **GCMD** Global Change Master Directory
- **GOSIC** Global Observing System Information Center
- **NDG** NERC Data Grid
- **NSDI** National Spatial Data Infrastructure
- **NSDL** National STEME Digital Library
- **NMMR** NOAA Metadata Manager Repository
- **OAI** Open Archives Initiative
- **SWEET** Semantic Web for Earth and Environmental Terminology
- **THREDDS** Thematic Realtime Environmental Data Distributed Services



## THREDDS Data Providers

- **University of Alabama Huntsville** (Sara Graves, Rahul Ramachandran, Steve Tanner, Ken Keiser)
- **ARM** (Atmospheric Radiation Measurement, Chris Klaus)
- **CDC**, the Climate Diagnostic Center (Roland Schweitzer)
- **COLA**, Center for Oceans Land Atmosphere (Joe Wielgosz)
- **University of Florence** (Stefano Nativi)
- **GMU**, George Mason University (Menas Kafatos and Ruixin Yang)
- **IRI/LDEO**, International Research Institute/Lamont Doherty Earth Observatory (Benno Blumenthal)
- **ESG**, the Earth System GRID (Luca Cinquini, NCAR/SCD)
- **IRIS DMC**, Incorporated Research Institutes for Seismology Data Management Center (Rob Casey)
- **NCAR**, the National Center for Atmospheric Research (Don Middleton)
- **NCDC**, the National Climatic Data Center (Ben Watkins)
- **NGDC**, National Geophysical Data Center (Ted Habermann)
- **NOMADS**, NOAA Operational Model Archive and Distribution System, (Glenn Rutledge, NCDC)
- **University of Oklahoma** (Kelvin Droegemeier)
- **PMEL**, the Pacific Marine Environment Laboratory (Steve Hankin)
- **FNMOC**, Fleet Numerical Meteorological and Oceanographic Center (Phil Sharfstein)
- **SSEC**, the Space Science and Engineering Center., U. of Wisconsin-Madison (Steve Ackerman, Tom Whittaker)
- **Unidata Community** ADDE servers (Tom Yoksas, Unidata Program Center)
- **CIESIN** (Consortium for International Earth Science Information Network, Bob Downs)
- **CUAHSI** (Consortium of Universities for Advancement of Hydrologic Science, David Maidment)
- **ESIG / NCAR** (NCAR Environmental Societal Impacts Group, Bob Harriss)
- **Earthscope** (UCAR UNAVCO, Chuck Meertens)
- **GEON** (GEOphysical Network, Chaitan Baru, UCSD San Diego Supercomputer Center)
- **ESRI GIS Community** (ESRI, Inc., Jack Dangermond, President)



## THREDDS Collaborators

- **ADDE**, Abstract Data Distribution Environment (University of Wisconsin – Madison, Tom Yoksas)
- **DIMES**, Distributed METadata System (George Mason University, Ruixin Yang)
- **DODS/OPeNDAP/Aggregation Server**, Distributed Oceanographic Data System/Open source Project for a Network Data Access Protocol (University of Rhode Island, Unidata, Ethan Davis)
- **DLESE**, Digital Library for Earth System Education (Rajul Pandya)
- **ESML**, Earth System Markup Language (University of Alabama-Huntsville, Rahul Ramachandran)
- **ESRI**, Environmental Science Research Institute (various)
- **GCMD**, Global Change Master Directory (Gene Major)
- **OGC and ISO Standards** (University of Florence, Stefano Nativi)
- **ADL** (Gazetteer Services The University of California, Santa Barbara, Linda Hill and Michael Goodchild)
- **DLESE Evaluation Services** (The University of Colorado CIRES, Susan Buhr)
- **DLESE Data Services** (Tamara Ledley)
- **DLESE Program Center** Digital Library for Earth System Education (Mary Marlino)
- **ESRI** (Jack Dangermond, President)
- **OPeNDAP** (The University of Rhode Island Open source Project for a Network Data Access Protocol -- formerly DODS, Peter Cornillon)
- **LAITS** (Laboratory for Advanced Information Technology and Standards, Liping Di, George Mason University)
- **NSDL Evaluation Services** (University of Colorado, Tamara Sumner)
- **OGC** (Open GIS Consortium, David Schell, President)
- **SWEET** (Semantic Web for Earth and Environmental Terminology, Rob Raskin)

# GCMD DODS/OPeNDAP Portal



A Global Change Master Directory Portal for the  
Open-source Project for a  
Network Data Access Protocol

### Find DODS/OPeNDAP-Related Data Sets by Topic:

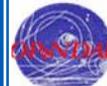
 <b>Agriculture</b> forestry, soils ...	 <b>Land Surface</b> land cover, erosion ...	Data Set Text Search <input type="text"/> <input type="button" value="Go"/> <a href="#">Advanced Search</a>
 <b>Atmosphere</b> precipitation, temperature ...	 <b>Oceans</b> circulation, salinity ...	
 <b>Biosphere</b> vegetation, zoology ...	 <b>Paleoclimate</b> tree rings, land records ...	
 <b>Climate Indicators</b> air temperature, drought ...	 <b>Snow and Ice</b> frost, snow cover ...	
 <b>Human Dimensions</b> environmental impacts, land use ...	 <b>Spectral / Engineering</b> radar, visible imagery ...	
 <b>Hydrosphere</b> snow & ice, water quality ...	 <b>Data Centers - Locations - Instruments - Platforms - Projects</b>	

 [Search the entire GCMD database](#)

---

 NASA Website Privacy and Security Statement, Disclaimer and Accessibility Certification  
Responsible NASA Official: Lola Olsen · Curator/Content Owner: Gene Major  
Contact GCMD User Support for assistance

<http://gcmd.gsfc.nasa.gov/Data/portals/dods/>



A Global Change Master Directory Portal for the  
Open-source Project for a  
Network Data Access Protocol

Keywords	Spatial Extent	Time Range
<p>This field is required.</p> <input type="text"/> in <input type="text" value="Full Text"/> <input type="radio"/> and <input type="radio"/> or <input type="text"/> in <input type="text" value="Full Text"/>	<p>This field is optional. Include? <input type="radio"/> YES <input checked="" type="radio"/> NO</p> <p>N [ ] W [ ]  E [ ] S [ ]</p> <p><a href="#">Click For Map</a></p> <p><a href="#">Trouble with the map applet?</a></p>	<p>This field is optional. Include? <input type="radio"/> YES <input checked="" type="radio"/> NO</p> <p>Time period of interest is during <input type="text"/> Jan. [ ] 1 1950 through Mar. [ ] 09 2004</p>
Select the number of records to view: <input type="text" value="15"/>		<input type="button" value="SEARCH"/> <input type="button" value="Reset"/>
Select the viewing format: <input type="text" value="Display"/>		<a href="#">Instructions</a>

 [Search the entire GCMD Database](#)

---

 NASA Website Privacy and Security Statement, Disclaimer and Accessibility Certification  
Responsible NASA Official: Lola Olsen · Curator/Content Owner: Gene Major  
Contact GCMD User Support for assistance

[http://gcmd.gsfc.nasa.gov/Data/portals/dods/freetext/ft\\_search.html](http://gcmd.gsfc.nasa.gov/Data/portals/dods/freetext/ft_search.html)

Selected Datasets

- ? SeaWinds on QuikSCAT Level 2B data product
- ? BRF - Black Rock Forest Open Lowland Terrestrial Station
- ? CDC\_NCEP\_Reanalysis\_Products\_Spectral
- ? CDC\_NCEP\_Reanalysis\_Products\_Surface\_Flux
- ? CDC\_NCEP\_Reanalysis\_Products\_Tropopause\_Level
- ? COADS 1-degree Equatorial Enhanced
- ? 5.5km East Coast AVHRR Archive

Additional Criteria

- Directory
- 1995 ...
  - 1996 ...
  - 1997 ...
  - 1998 ...
  - 1999 ...
  - 2000 ...
  - 2001
    - 1 ...
    - 10 ...
    - 11
    - 12 ...
    - 2 ...
    - 3 ...
    - 4 ...
    - 5 ...
    - 6 ...
    - 7 ...

- 195 files:
- d01305070339.pvu.Z
  - d01305165022.pvu.Z
  - d01305182858.pvu.Z
  - d01306065308.pvu.Z
  - d01306181823.pvu.Z
  - d01307064240.pvu.Z
  - d01307180747.pvu.Z
  - d01308063205.pvu.Z
  - d01308175700.pvu.Z
  - d01309062233.pvu.Z
  - d01309174621.pvu.Z
  - d01310061107.pvu.Z
  - d01310174031.pvu.Z
  - d01311060044.pvu.Z
  - d01311172542.pvu.Z
  - d01312055023.pvu.Z
  - d01312072950.pvu.Z
  - d01312171523.pvu.Z

Directory Regex

- lat: lat=[0 : 1023]
- lon: lon=[0 : 1023]
- Grid {
  - dsp\_band\_1: lat=[350 : 351]lon=[500 : 512]
  - lat: lat=[350 : 351]
  - lon: lon=[500 : 512]

Remove Selected Remove All Favorite

Select All Deselect All Reset CE

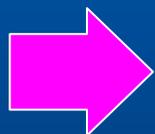
Output Settings

Output

Send to View Panel Send to File Send to Standard Out

The ODC

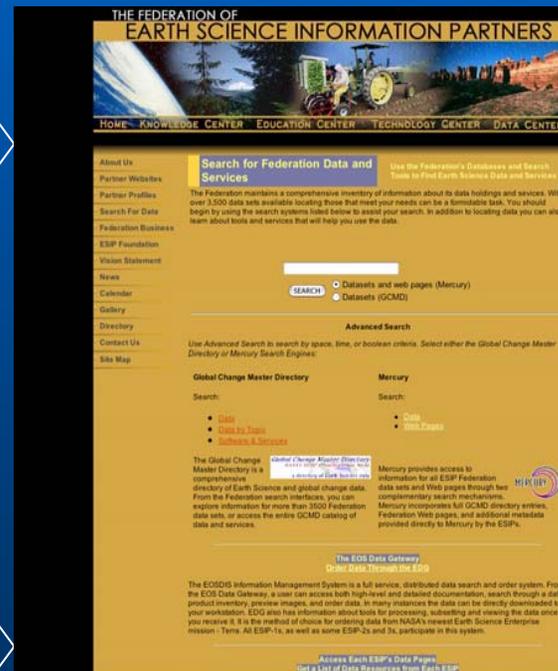
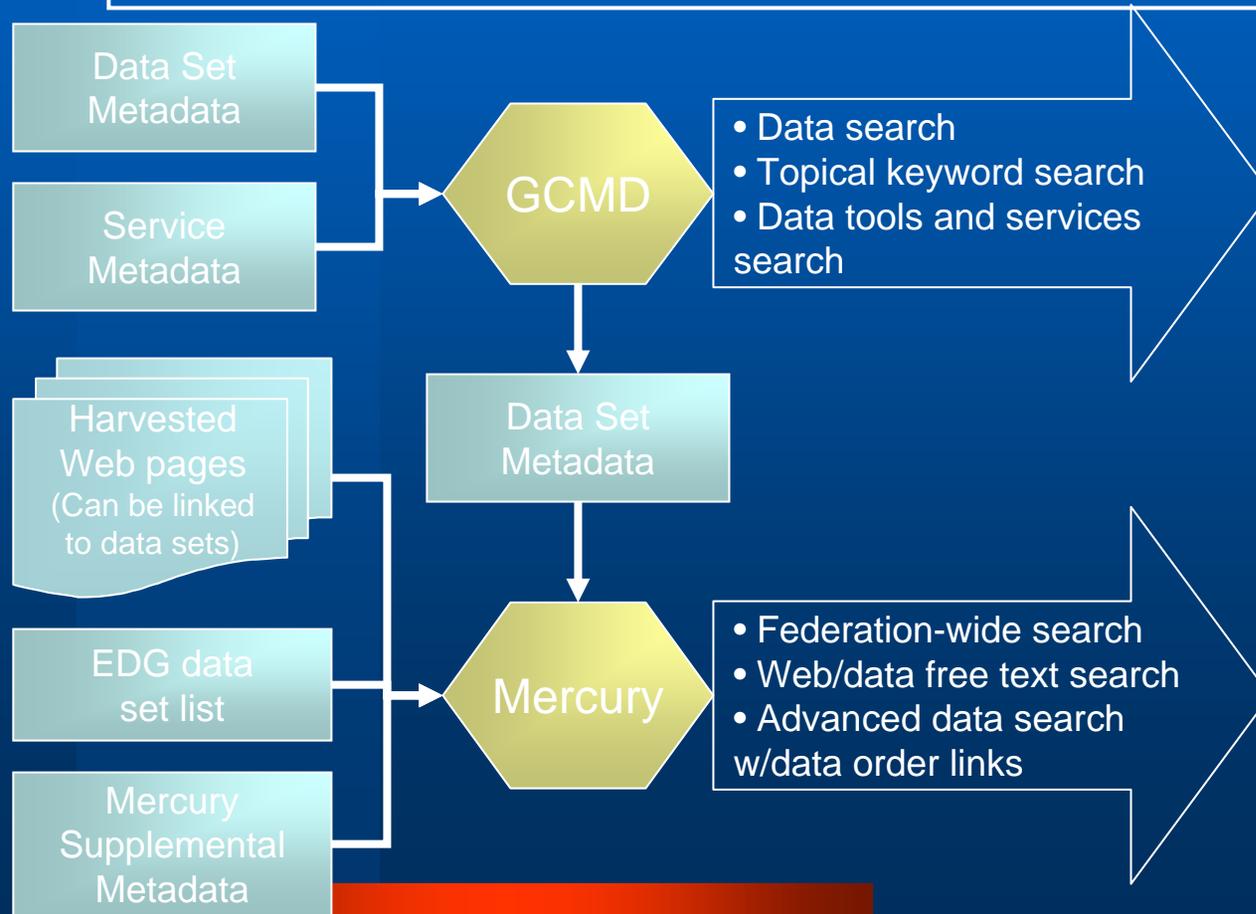
- Collaborations require long-term maintenance of both the data and descriptions of the data; i.e., metadata.



*The degree of system interoperability is determined by the associated metadata and the quality of that metadata.*

# Federation Interactive Network for Discovery

FIND combines the capabilities of two search systems (the Global Change Master Directory and Mercury). It provides users with a rich set of options to locate ESIP Federation data, services, and information. FIND is accessible from the Federation Home Page.

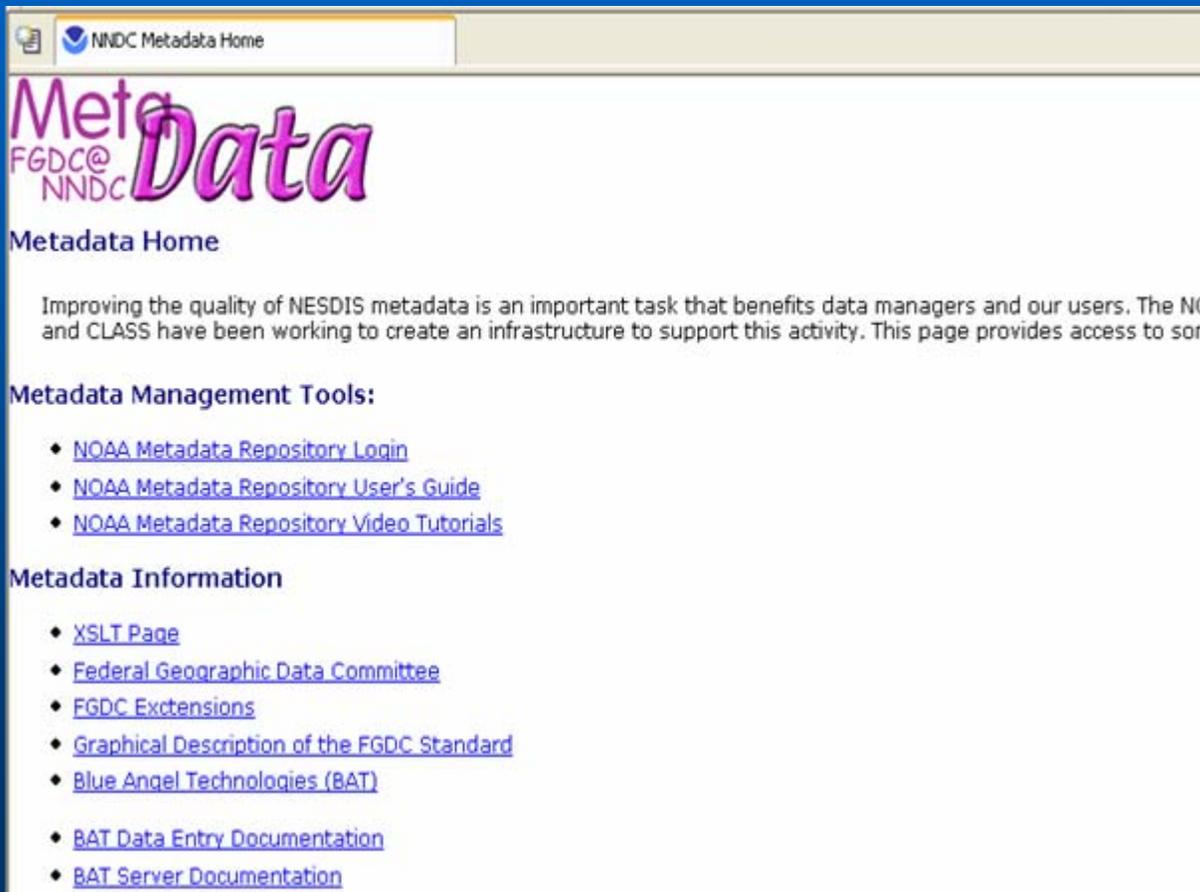


<http://www.esipfed.org/find>

- NESDIS Metadata Working Group
  - A good first start.
  - Community wide audience needed.

How?  
GeoScience Technology Forum (GTF)

- NSF Cyber Infrastructure
- GEO ??



The screenshot shows a web browser window titled "NNDC Metadata Home". The page features a large "Meta Data" logo in purple and pink, with "FGDC@ NNDC" written below it. The main heading is "Metadata Home". Below this, a paragraph states: "Improving the quality of NESDIS metadata is an important task that benefits data managers and our users. The NNDC and CLASS have been working to create an infrastructure to support this activity. This page provides access to some of the tools and information available." The page is organized into two main sections: "Metadata Management Tools:" and "Metadata Information:". Under "Metadata Management Tools:", there are three bullet points with links: "NOAA Metadata Repository Login", "NOAA Metadata Repository User's Guide", and "NOAA Metadata Repository Video Tutorials". Under "Metadata Information:", there are six bullet points with links: "XSLT Page", "Federal Geographic Data Committee", "FGDC Extensions", "Graphical Description of the FGDC Standard", "Blue Angel Technologies (BAT)", "BAT Data Entry Documentation", and "BAT Server Documentation".

## *NOAA and other Programs*

- NOAA's Scientific Data Stewardship (SDS) well conceived.
- CLASS requires more community involvement and they are actively seeking feedback. The time is now to design interoperability into CLASS. Re-engineering difficult.
- Many efforts now exist from which to leverage:  
GCOS, IOOS, US Oceans, DMAC, NVOADS, WCRP, IPCC, WMO, GCMD, THREDDS, (more on this)...
- NOAA's Office of Project Planning and Implementation now formed.



- Affirmed need for timely, quality, long-term, global information as a basis for sound decision making.
- Recognized need to support:
  - 1) **Comprehensive, coordinated, and sustained** Earth observation system or systems;
  - 2) Coordinated effort to **address capacity-building** needs related to Earth observations;
  - 3) **Exchange of observations in a full and open** manner with minimum time delay and minimum cost; and
  - 4) Preparation of a **10-year Implementation Plan**, building on existing systems and initiatives by European ministerial in late 2004
- Established *ad hoc* Group on Earth Observations (GEO) to develop Plan
- Invited other governments to join.

- A system of systems can be designed with active involvement with existing data managers system managers and scientists: ESIP Role?
- Leveraging intra-Agency activities with GEO 10-year plan as the driver.

How is our community addressing the needs of GEO?

- To overcome a deficiency in model data access, some of the Nations top scientists are actively engaged in a grass-roots framework to share data and research findings over the Internet.
- NCDC, NCEP and GFDL initiated the NOAA Operational Model Archive and Distribution System.
- NOMADS is a distributed data services pilot for format independent access to climate and weather models and data.

**NOMADS**  
 The NOAA Operational Model  
 Archive and Distribution System

**GO-ESSP**

**Core NOAA NOMADS Collaborators**

- ◆ Climate Diagnostics Center (CDC) Boulder, CO
- ◆ Geophysical Fluid Dynamics Laboratory (GFDL) Princeton, NJ
- ◆ National Climatic Data Center (NCDC) Asheville, NC (Project Lead)
- ◆ National Centers for Environmental Prediction (NCEP) Camp Springs, MD
- ◆ Pacific Marine Environmental Laboratory (PMEL) Seattle, WA
- ◆ NOAA Forecast Systems Laboratory (FSL) Boulder, CO

**External Core Collaborators**

- Center for Ocean-Land-Atmosphere Studies (COLA) (Maryland)
- Department of Energy's Argonne, Los Alamos, Oak Ridge, Lawrence Berkley, Livermore National Laboratories & Information Sciences Institute (ISI), University of Southern California under the Earth System Grid Project
- National Center for Atmospheric Research (NCAR) Colorado
- Unidata Program Center (UCAR/Unidata) Colorado
- LLNL Program for Climate Model Diagnosis and Intercomparison
- NASA's Global Change Master Directory (GCMD) Maryland
- National Coastal Data Development Center
- University of Rhode Island (OPeNDAP Consortium)

**External Collaborators include**

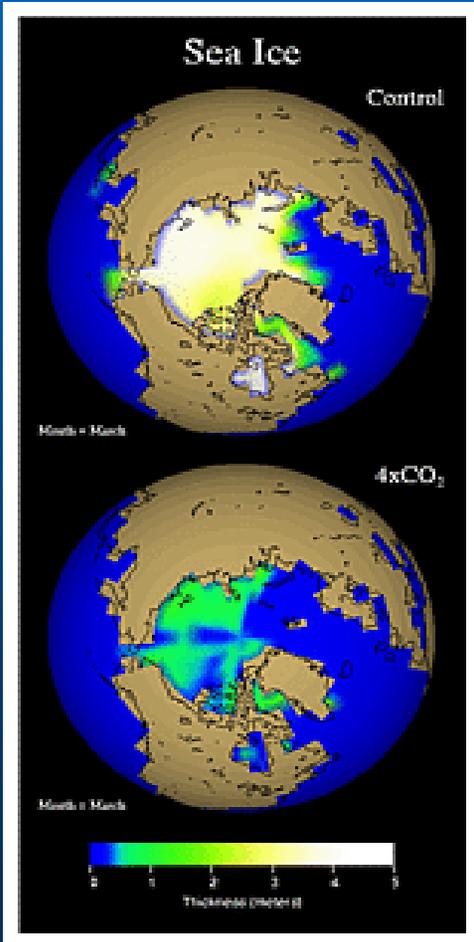
- Center for Earth Observing and Space Research (CEOSR), NASA-GSFC Maryland
- George Mason University (NASA SI-ESIP), Virginia
- National Severe Storms Laboratory (NSSL), Oklahoma/SSEC University of Wisconsin
- Universities of Alabama (Huntsville), California (Santa Barbara), Washington & Iowa St.
- National Science Foundation (NSF) CyberInfrastructure

**International Participants**

- British Atmospheric Data Center, Oxfordshire, United Kingdom)
- UK's Natural Environment Research Council (NERK DataGrid Project)
- Committee for Earth Observing Satellites (CEOS) Grid Project
- Climate Action Partnership (CAP), BOM Australia (US Depts. of Commerce, Energy, State, and EPA)

Logos for participating organizations: Department of Commerce, NOAA, University of Colorado, NASA, Unidata, and NCAR.

- provide distributed access to models and associated data,
- promote model evaluation and product development,
- foster research within the geo-science communities (ocean, weather, and climate) to study multiple earth systems using collections of distributed data,
- develop institutional partnerships via distributed open technologies.



- The users experience is often frustrating—
  - What data of interest exist?
  - Are they going to be useful to me?
  - How can I obtain them in a usable form?
- Time and effort are wasted on data access and format issues.
- As a result atmosphere/ocean/climate data are under-utilized. Model inter-comparison nearly impossible.



## NOMADS

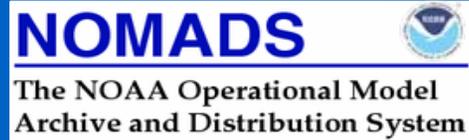
The NOAA Operational Model  
Archive and Distribution System



## *Scientific Data Networking...*

NOMADS simplifies scientific data networking, allowing simple access to high volume remote data, unifying access to Climate and Weather models:

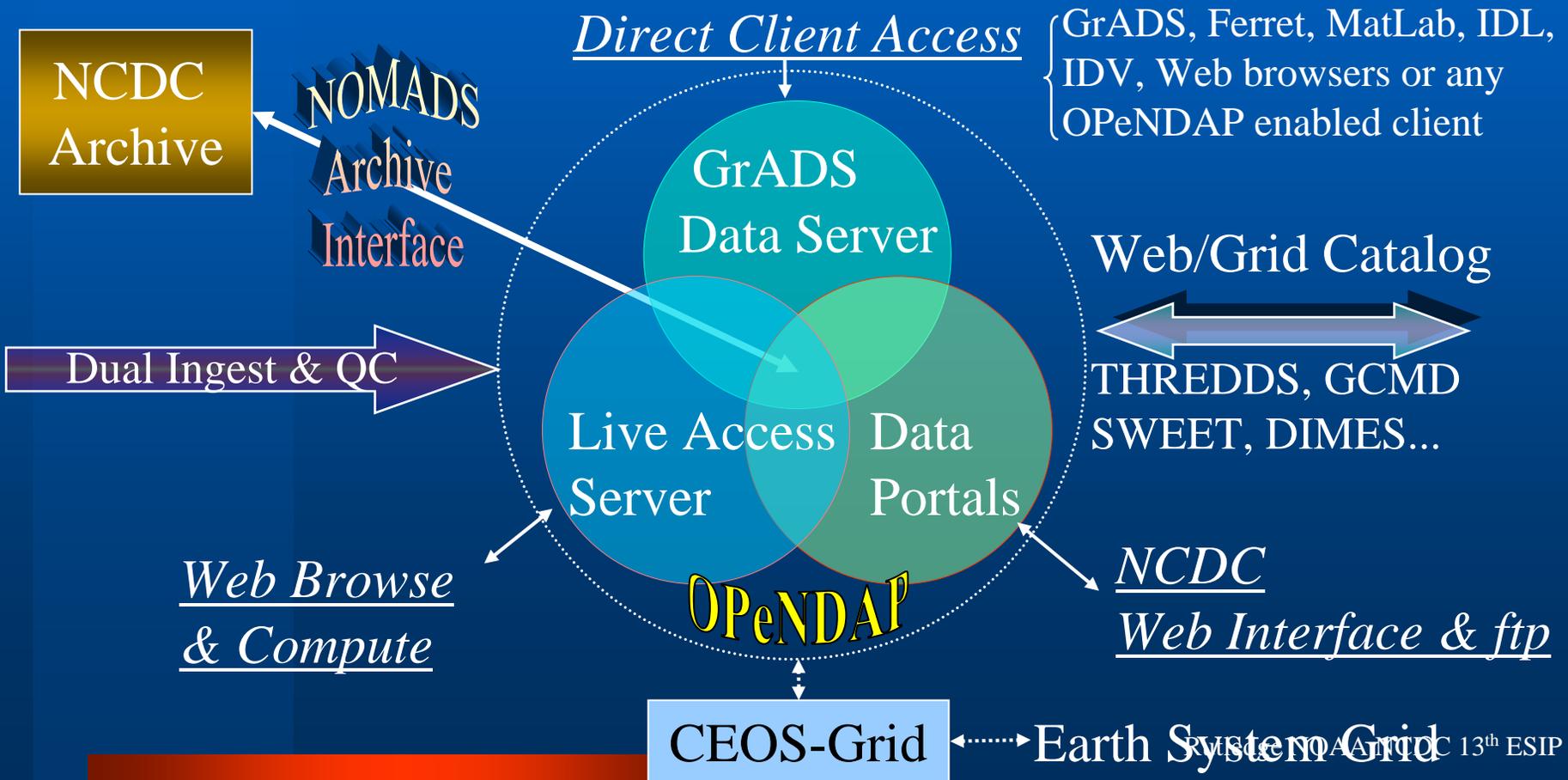
- **Data access (client)**
  - Access to remote data in the users normal application
    - IDL / IDV / Matlab / Ferret
    - GrADS (GRIB/BUFR w/ GDS)
    - Netscape / Excel / http (wget)
    - CDAT (PCMDI)
    - **Any netCDF application** (i.e., AWIPS)
  - **Don't need to know the format in which the data are stored.**
- **Data publishing (server)**
  - **Can serve data in various formats**
    - netCDF / GRIB / BUFR / GRIB2
    - HDF (3-5) / EOS
    - SQL / FreeForm
    - JGOFS / NcML
    - DSP
    - ascii, others...
  - **Spatial and temporal sub-setting and host side computations on the fly.**



## *Collaborating Programs*

CAP	Climate Action Partnership	DOC DOE EPA State Dept
CDP	Community Data Portal	NCAR
CEOS	Committee on EO Satellites	NOAA Representative
CEOP	Coordinated Earth Obs Period	NOAA Representative
EPA	Air Quality Models	(in progress)
GO-ESSP	Earth Science Portal	Founding Member
NASA GCMD		Science Advisory Board
NERC DataGrid		Advisory Committee
NSF Cyberinfrastructure		Member
NSF LEAD & Geo-Science Tech Forum (GTF)		Data / Planning Committee
NVODS / US GODAE / GOOS		Data Provider
Unidata THREDDS, NSDL, DLESSE		Data Provider
WCRP World Climate Research Program		JSC/CLIVAR Briefings

### Multiple paths to format independent data access:





- NOMADS uses the Open Source http based OPeNDAP.
- OPeNDAP is a binary-level protocol designed for the transport of scientific data subsets over the Internet. Provides server side data manipulation on-the-fly (e.g., GrADS-DODS).
- Data formats: GRIB, GRIB2, BUFR, HDF, NetCDF, ascii...  
Conventions: COARDS, CF, FGDC, DIF...libraries built as necessary.
- APIs: JAVA-OPeNDAP, C++-OPeNDAP, NetCDF, GRIB, BUFR, THREDDS, Python.



- A grass roots effort has formed by data managers called the **Global Organization for Earth Systems Science Portals**  
**GO-ESSP** <http://esportal.gfdl.noaa.gov>

- Unidata
- ESG (NCAR, LLNL)
- OPeNDAP
- COLA
- NOMADS (GFDL, PMEL, NCDC, NCEP, others)
- NASA/GCMD
- BADC, BODC
- WMO



## GO-ESSP

- The Global Organization for Earth System Science Portal (GO-ESSP) is a collaboration designed to build the infrastructure needed to create web portals to provide access to observed and simulated data within the climate and weather communities.
- The infrastructure created within GO-ESSP will provide a flexible framework that will allow interoperability between front-end and back-end software components. GO-ESSP is an international collaboration involving software developers from both Europe and the United States.

## *Data Availability Overview*



**CDC:** Reanalysis, climate weather models, in-situ

**GFDL:** Coupled Models, Control and Perturbation Integrations and historical 20th century simulations using solar, volcano, GHG and aerosol forcings.

**FSL:** MADIS mesoNets, Hi-Res RUC-II

**NCAR:** Community Climate System Model / Land Surface CO2 predictive models (VEMAP), Reanalysis / Eta

**NCDC:** Archive for NCEP model input/output / Select NCDC Observation datasets, Ocean/Ice WAVE, NARR, SST's...

**NCEP:** Real-time Input/Output, Reanalysis (I&II), Ensembles, Sea Ice Ocean, CDAS, Hourly Eta, Climate Forecast Models...

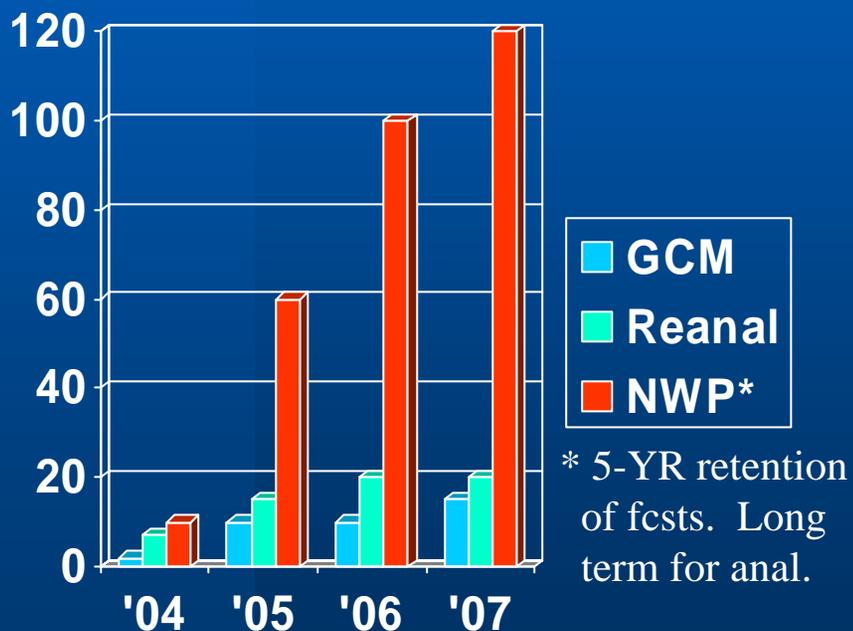
**LLNL:** AMIP / Probabilistic information

**PMEL:** Ocean and Climate datasets

- **NCDC NOMADS Archive**
  - POR: 2002 to Real-Time
  - **Eta** (12km); **GFS** (1 degree); **GDAS**; **NARR** 12km 30yrs
  - RUC-II 20/40km; Ocean and Ice WAVE Models
  - NCDC Reference Data Sets (Reynolds SST's, GHCN...)
  - NCDC Mirror site to NCEP NOMADS for Eta & GFS
- **NCEP Real-Time NOMADS**
  - Global Forecast System GFS 1 degree
  - **Hourly Eta** at 12km
  - Regional Spectral Model (RSM) and **Ensembles**
  - Climate Data Assimilation System (CDAS)
  - AMIP Climate Monitoring, **Climate Forecast Model**
  - NCEP/NCAR Global Reanalysis 1&2

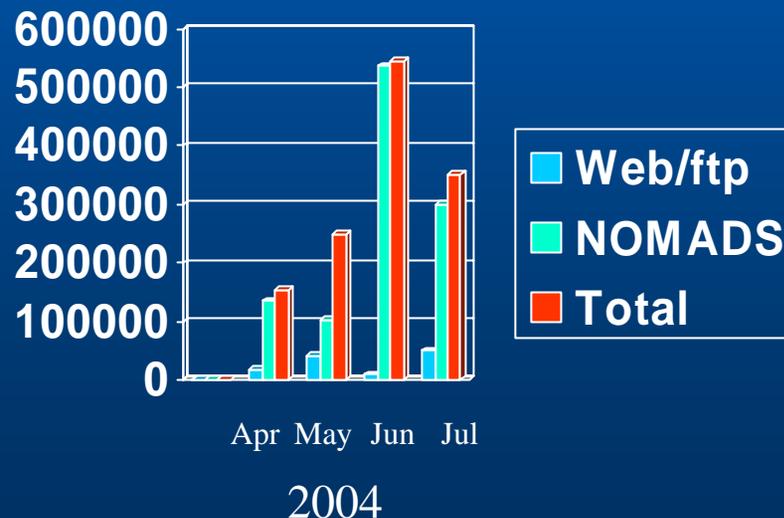
- **Data Philosophy and Retention**
  - Data are free.
  - NWP forecast data are retained for five years.
  - Analysis, Reanalysis, observations, and GDAS model input are retained for long term stewardship.
- **Data Users**
  - Resolution of IP addresses indicate a broad range, and consistent use of NOMADS available data:
    - U.S. Agencies, Academic Institutions: K-12 to Research
    - International governments, (Italy, Japan, countries within South America and Africa. Many others).
    - Private Sector and Non-Government Organizations NGO's
    - World Bank, United Nations (FAO), others.

## NCDC Ingest Volume Tb/Yr



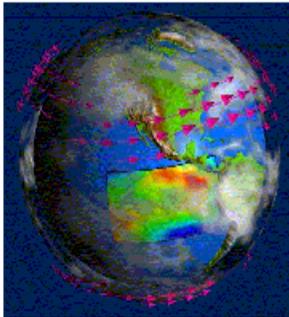
Existing and Projected Volume

## NOMADS Downloads / month



[NCDC](#) | [Contents](#) | [Satellite](#) | [Climate](#) | [Radar](#) | [Model](#) | [Search](#) | [Help](#)  
[Model Resources](#) | [About](#) | [Inventories](#) | [Get/View Data](#) | [Publications](#) | [Other Resources](#)

### Model Resources



- ◆ [About NCDC's Model Resources](#)
- ◆ [Model Data Inventories](#)
- ◆ [Get / View Model Data](#)
- ◆ [Publications](#)
- ◆ [Other Model Data Resources](#)

[NCDC](#) | [Contents](#) | [Satellite](#) | [Climate](#) | [Radar](#) | [Model](#) | [Search](#) | [Help](#)  
[Model Resources](#) | [About](#) | [Inventories](#) | [Get/View Data](#) | [Publications](#) | [Other Resources](#)

Three primary methods for data access:

- Web Interface
- OPeNDAP
- ftp w/ on the fly Grib subsetting

On-line or  
Off-line (archive)

- Server-side data computations...

```

* January Mean 500 Height (1981 to 1989) minus (1990 to 1998)
* Mean & Standard Deviation for all 10 ensembles
* Time required: 60 secs
'reinit'
'!date'
* baseURL = 'http://motherlode.ucar.edu:9090/dods/_expr_'
* GKR 2/13/03 New NCAR URL
baseURL = 'http://dataportal.ucar.edu:9191/dods/'
expr = 'ave(z,t=387,t=483,12)-ave(z,t=495,t=591,12)'
xdim = '0:360'
ydim = '20:90'
zdim = '500:500'
tdim = '1nov1978:1nov1978'

```

At left is the complete script for generating mean and sdev at 500mb analyzing 18 years of “Climate of the 20<sup>th</sup> Century” over the Internet:

Traditional vs. NOMADS methods:

Data volume transported: 100Gb vs. 2Kb

Time to access data: 2 days vs. 60 sec

Code development: days vs. minutes

Fortran based LOC: 1000 vs. 50 LOC

```

'define resa = result.1'
'define resb = result.2'
'define resc = result.3'
'define resd = result.4'
'define rese = result.5'
'define resf = result.6'
'define resg = result.7'
'define resh = result.8'
'define resi = result.9'
'define resj = result.10'
say 'got data'

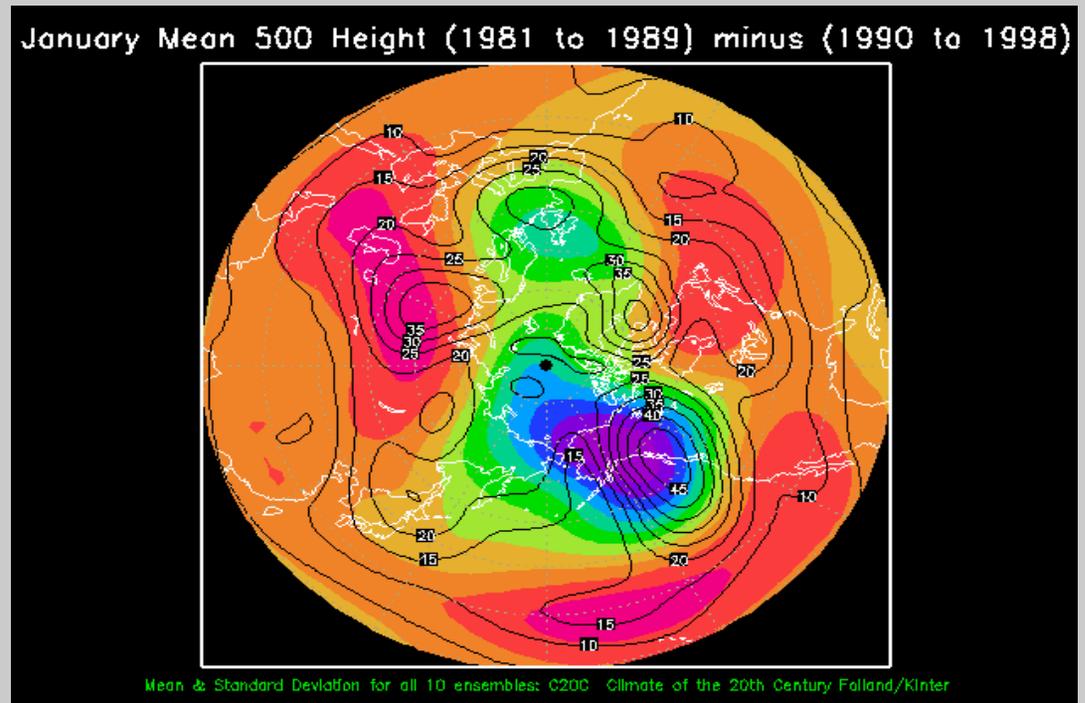
'set lev 500'
'set lat 20 90'

'define mean = (resa + resb + resc + resd + rese + resf + resg + resh + resi + resj)/10'

'define d1 = (pow(resa-mean,2)) ; 'define d2 = (pow(resb-mean,2))'
'define d3 = (pow(resc-mean,2)) ; 'define d4 = (pow(resd-mean,2))'
'define d5 = (pow(rese-mean,2)) ; 'define d6 = (pow(resf-mean,2))'
'define d7 = (pow(resg-mean,2)) ; 'define d8 = (pow(resj-mean,2))'
'define d9 = (pow(resi-mean,2)) ; 'define d10 = (pow(resj-mean,2))'
'define stddev = pow((d1 + d2 + d3 + d4 + d5 + d6 + d7 + d8 + d9 + d10)/10,0.5)

'set gxout shaded'
'set mproj nps'
'display mean'
'draw title January Mean 500 Height (1981 to 1989) minus (1990 to 1998)'
'set string 3 bc 1'
'draw string 5.5 .5 Mean & Standard Deviation for all 10 ensembles:
'C20C Climate of the 20th Century Folland/Kinter'
*cbarh'
'set gxout contour'
'set ccolor 0'
'display stddev'
'!date'

```



### The NCDC Web Interface originally developed at NCEP:

NCEP NWP Model Datasets in NCDC Repository

model	grid/scale	freq	plot	ftp	http	nomads gds	contact 1	contact 2
<b>GFS Analysis and Forecasts</b>								
GFS-AVN	<a href="#">201</a>	00,06,12,18Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-AVN	<a href="#">202</a>	00,06,12,18Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-AVN	<a href="#">203</a>	00,06,12,18Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-AVN	<a href="#">211</a>	00,06,12,18Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-AVN	<a href="#">213</a>	00,06,12,18Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-MRF	<a href="#">201</a>	00Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-MRF	<a href="#">202</a>	00Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-MRF	<a href="#">203</a>	00Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
GFS-MRF	<a href="#">205</a>	00Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
<b>ETA Analysis and Forecasts</b>								
Early-ETA	<a href="#">212</a>	00,12Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
Meso-ETA	<a href="#">211</a>	00,12Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>
Meso-ETA	<a href="#">212</a>	00,06,12,18Z	<a href="#">plot</a>	<a href="#">ftp4u</a>	<a href="#">http</a>	<a href="#">gds</a>	<a href="#">Glenn Rutledge</a>	<a href="#">T</a>

Variable:

- capes 1 level \* Convective Available Potential Energy (Surface) [J/kg]
- cins 1 level \* Convective Inhibition (Surface) [J/kg]
- p 1 level \* Total Precipitation [kg/m^2]
- pc 1 level \* Convective Precipitation [kg/m^2]
- ps 1 level \* Surface Pressure [Pa]
- pwat 1 level \* Entire Atmosphere Precipitation [kg/m^2]
- rh2m 1 level \* Meter Relative Humidity [%]
- slpe 1 level \* Sea Level Pressure, ETA re...
- t2m 1 level \* Meter Temperature [K]
- u10m 1 level \* Meter U Winds [m/s]
- v10m 1 level \* Meter V Winds [m/s]

Note: Some of the above listed variables may not be available at the time or model level. To see what data is present, use your Back button to return to the main page and use the review links at the bottom of the page or see the variable list.

Level: 1

extra operation 1: (none) 2: (none)

Data available from 00Z 28 dec 2003 to 06Z 29 dec 2003 at 6 hour intervals

Time 00Z 28 dec 2003

Map projection: lat-lon (180E) only for custom maps long: 280 long width: 50 lat: -60 lat height: 60

Draw: shaded Contour interval: def white: def Plot size: 800x600

NOMADS leverages efforts across the community.

- NCDC NOMADS ingests 150K grids day. POR 2002 to present.

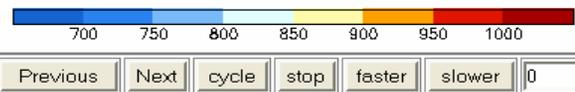
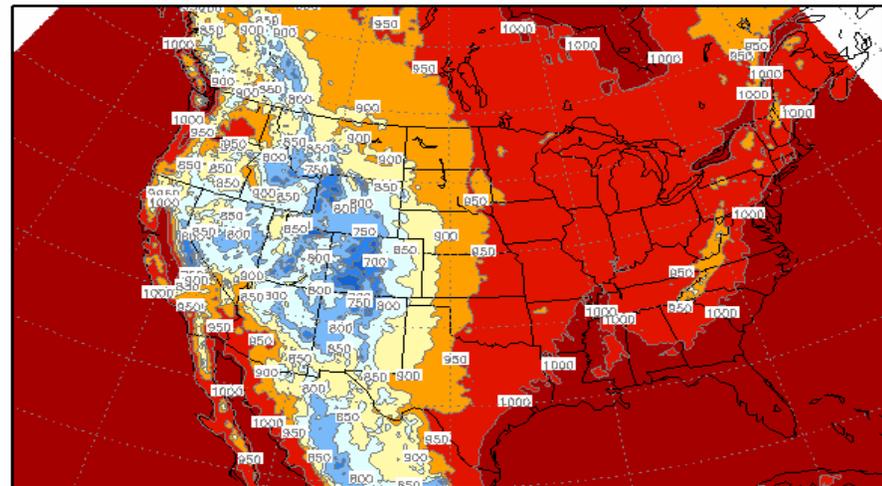
- Any one of these accessible in seconds  
Via: OpENDAP  
GDS  
ftp  
Web Plotter  
LAS (soon)

NOMADS Interactive Web Plotter - Order # 1052 ./meso-eta-hi\_218\_20040529\_1800\_fff.ctd

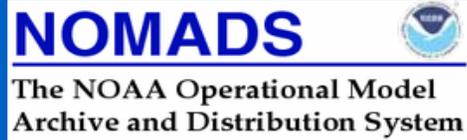
PRESsfc 1000

18Z29may2004 to 06Z01jun2004

PRESsfc 18Z29MAY2004

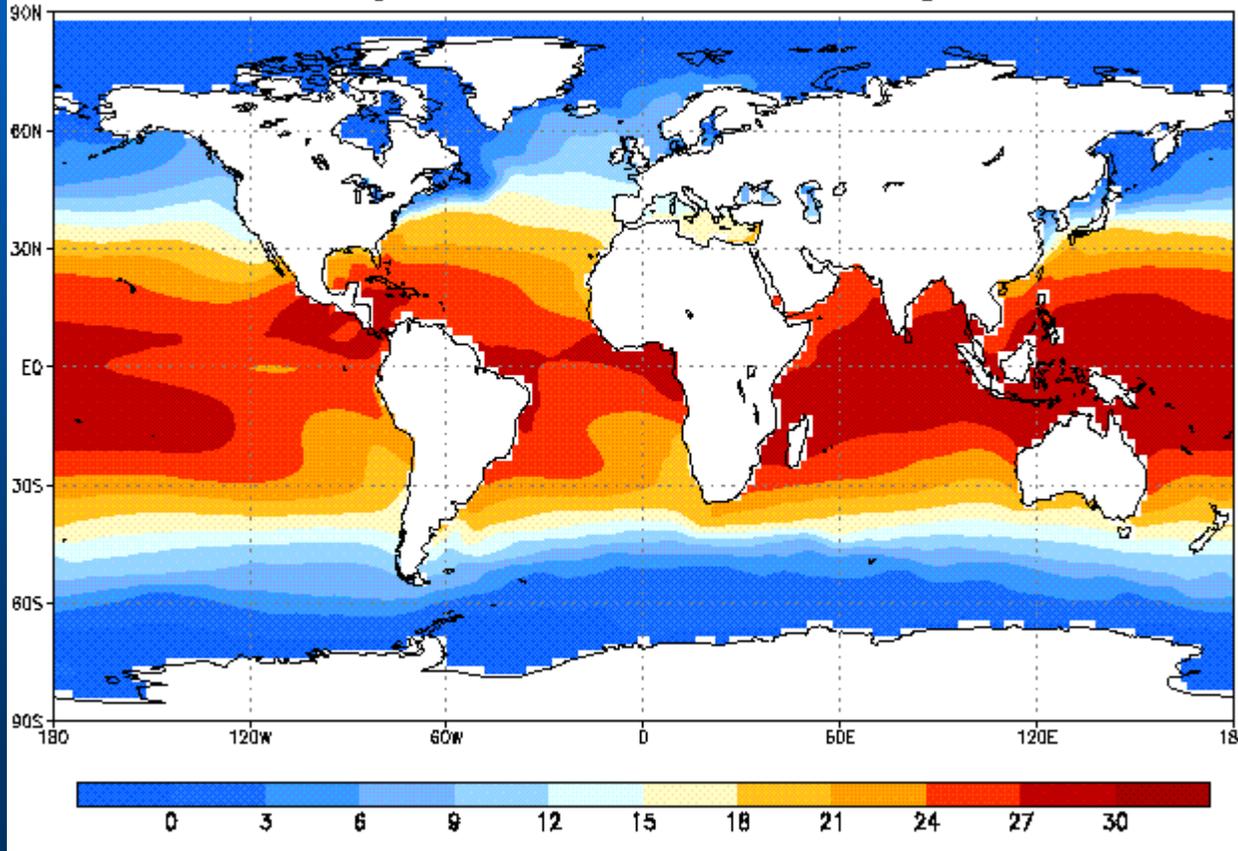


# Promoting Model to Obs. Intercomparisons



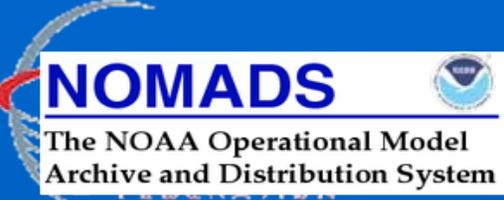
## *NCDC Reference Datasets*

Smith Reynolds ERSST Jan 1854 Degrees C



NCDC reference  
And others  
datasets also available:

- CARDS (IGRA)
- GHCN
- NARR
- Ocean WAVE

The logo for NOMADS (The NOAA Operational Model Archive and Distribution System) features a stylized globe with a red and blue arc on the left side. The text "NOMADS" is written in a bold, blue, sans-serif font.

**NOMADS**

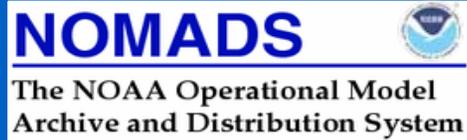
The NOAA Operational Model  
Archive and Distribution System

13<sup>th</sup> Federation Meeting

## National Digital Forecast Database

- Value added retailers who make value added products can use NOMADS GDS to get the meteorological data they need without downloading entire files.
- Users (forecasters) of NDFD can create their own products using GDS server accessing only data they need.
- GDS reduces the bandwidth needed to create products in weather service operations.
- For internet-2 bandwidth, servers at Regional Centers can distribute data to WFO's for their operations.

# Enabling private sector access: An example



## NOMADS Ensemble Access

TODAY is: 2004, 08, 09

Please select :

### NOMADS Ensemble Probabilities on the fly

STATION NAME

ASHEVILLE MUNICIPAL NC US

STATION Latitude: 35.43 STATION Longitude: -82.55

Date (HR/DD/MM/YY)GMT

14 09 08 04

Cycle

00z

Create an event:

### ● No need for image generation of ensembles...

Temperature:

Notes: you can create a temperature event by giving a lowest temperature or a highest temperature or a range of temperature. For example, for freezing event, giving lowest temperature lower than 32F and do not check the highest temperature.

Lowest TEMP:

Higher than

UNIT: K

Highest TEMP:

Higher than

UNIT: K

Precipitation

Higher than mm/day

Wind Speed

Higher than

UNIT: m/sec

Click YES to show URL query for ensemble members:

OPeNDAP constraint expression

URL is: [http://nomad3.ncep.noaa.gov:9090/dods/enshires/archive/ens20040809/ensc0\\_00z\\_1x1.ascii?pratesfc\[3:3\]\[125:125\]\[277:277\]](http://nomad3.ncep.noaa.gov:9090/dods/enshires/archive/ens20040809/ensc0_00z_1x1.ascii?pratesfc[3:3][125:125][277:277])

NO

YES

EventProbability

Reset

## *CLASS and NOMADS*

- Under NOAA's Scientific Data Stewardship (SDS) programs the NOAA Comprehensive Large Array-data Stewardship System (CLASS) will act as the main portal for NOAA/NESDIS environmental data, providing physical archive, access, and distribution capabilities for large array data sets.
- The NOMADS team and its collaborators are working with CLASS as the system progresses through its phased implementation plans for access to weather and climate models via OPeNDAP and OPeNDAP Servers (GDS/LAS). Metadata management must be addressed at the Agency level.

# Next Steps to Collaborations

#1

- *Leverage the resources as goals outlined by the Group on Earth Observations – GEO (& Earth Observation Summit) through appropriate Agency working groups and representatives:*

- Interagency Working Group on Earth Observations (IWGEO)  
Data and Information Systems (OWGDIS)

# Next Steps to Collaborations 2

#2

• *Ensure that NGO's, University, or Institutional partners are involved in this process e.g.,*

- COLA
- EOGEO
- many more...

# Next Steps to Collaborations 3

#3

- *Agencies partially fund (5%?) data management for each program. This should not be considered a separate activity.*

# Next Steps to Collaborations 4

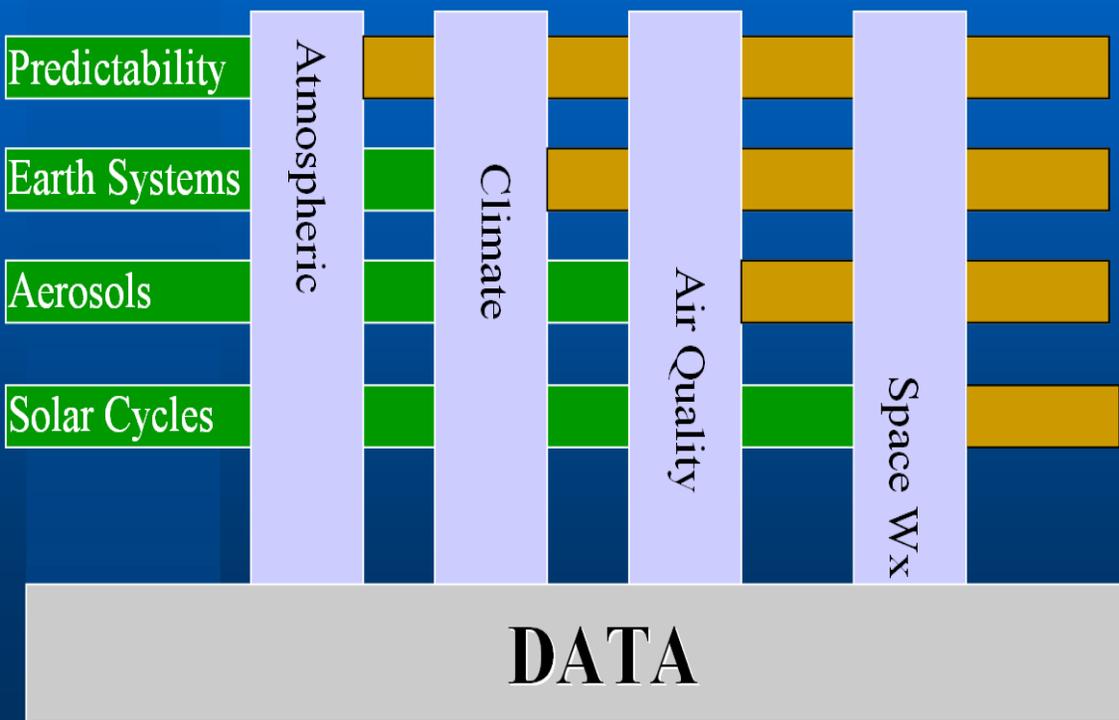
#4

- *Engage and leverage from existing efforts and organizations especially NASA, NOAA, NSF, etc.*
- NSF **CyberInfrastructure** (Ad Hoc Committee June 2004)
- GO-ESSP
- LEAD **GeoScience Technology Forum** (GTF)
- THREDDS / GCMD / SWEET / FIND / ...
- NERC Data Grid (Europe)
- WMO CBS

# Next Steps to Collaborations 5

#5

- Advance Agency Program Management at the Data level.



# Next Steps to Collaborations 6

#6

- Advance the building of Ontologies at Data Centers and Providers, (with SWEET), to interact with an enhanced “THREDDS & GCMD” effort for data search and access at the “variable” level\*.

\* Using OPeNDAP enabled clients and Servers

*For more information...*

- For NOMADS Program Information see:  
<http://www.ncdc.noaa.gov/oa/climate/nomads/nomads.html>
- For NOMADS Model Data Access:  
NOAA NCDC Main Page → Climate → *Model Resources*  
<http://nomads.ncdc.noaa.gov>
- Or contact:  
Glenn.Rutledge @ noaa.gov
- *Selected Publications on distributed data access and NOMADS:*  
<http://www.ncdc.noaa.gov/oa/model/publications/publications.html>

**QUESTIONS ?**