Responses to Recommendations from the May 2004 GCMD Science User Working Group (UWG) Report

I. Strategic Recommendations

1. Develop a Vision for the GCMD.

The Global Change Master Directory project envisions becoming the world's premiere resource to discover scientific data and provide direct access to selected data sets. Through the wise design and maintenance of internal applications, the GCMD also envisions enabling the discovery of and direct access to relevant data services for the analysis and visualization of data. Project activities are designed to promote the stewardship of scientific resources and thereby build the knowledge and resource connections to enhance our understanding of the Earth, the Solar System, and the Universe.

a) Coordinate work with other cyberinfrastructure efforts.

The GCMD is engaged in efforts towards interoperability and sharing of data (and metadata) across multi-discipline boundaries. Ongoing collaborations with the Marine Metadata Interoperability (MMI), Thematic Realtime Environmental Distributed Data Services (THREDDS), NASA's EOS ClearingHOuse (ECHO), and others are manifestations of our work within the cyberinfrastructure (CI) domain. At a recent (October 2005) gathering of CI professionals at the National Forum for Geosciences Information Technology (FGIT), the GCMD was recognized as being an active participant. See figure below.

Source: Wilhelmson, R. National Forum for Geosciences Information Technology Inaugural Meeting, Washington, D.C. October 2005.

b) Align with NASA and national objectives.

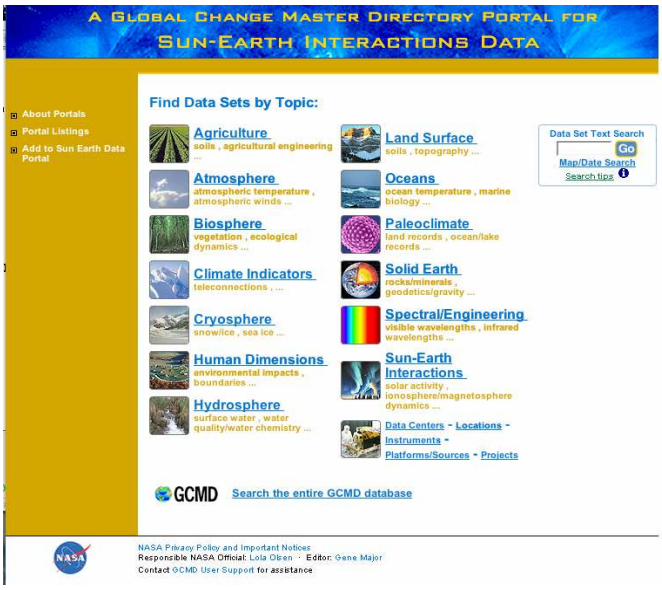
The Directory Interchange Format (DIF) for describing Earth Science data, and the Service Entry Resource Format (SERF), along with the software behind the search and retrieval, are remarkably flexible and can be adapted for describing lunar, planetary, and astronomical data and subsequently linking these to related services. Central to the GCMD are robust hierarchies of controlled Earth Science keywords for search and retrieval. Targeted controlled vocabularies can be created or adopted for use in searching for space science data sets - in keeping with NASA's strategic goals.

GCMD continues to assess the re-alignments of NASA. Possibilities include recombining astronomy and space science metadata with Earth science metadata. Some recent activities in this area include (See Section 3a):

- Added Space Science as a Category for future space science keywords. These were added to describe some Antarctic research data sets funded by the NSF Polar Programs. See Section 3.
- Expanded the Sun-Earth Interactions keywords by adding more Terms and Variables to describe NASA Earth-Sun missions:

EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Aurorae
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Electric Fields/Electric Currents
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Geomagnetic Forecasts
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Geomagnetic Indices
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Ion Chemistry/Ionization
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Magnetic Fields/Magnetic Currents
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Magnetic Storms
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Plasma Waves
EARTH SCIENCE > Sun-earth Interactions > Ionosphere/Magnetosphere Dynamics > Solar Wind
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Corona Holes
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Coronal Mass Ejections
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Coronal Properties
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Active Regions
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Flares
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Imagery
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Irradiance
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Oscillations
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Prominences/Solar Filaments
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Radio Wave Emissions
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Synoptic Maps
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Ultraviolet Emissions
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar Velocity Fields
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Solar X-ray Emissions
EARTH SCIENCE > Sun-earth Interactions > Solar Activity > Sunspots
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Alpha Particle Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Electron Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Heavy Nuclei Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Ion Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Neutral Particle Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Proton Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Flux > Sub-atomic Particle Flux
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Energy Deposition
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Particle Composition
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Particle Density
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Particle Distribution Functions
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Particle Speed
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Particle Temperature
EARTH SCIENCE > Sun-earth Interactions > Solar Energetic Particle Properties > Total Electron Content

- Mapped the GCMD keywords to the 12 National Applications and NASA Science focus areas and utilized that information to create application-oriented portals. See Section 2.
- Created a portal for Sun-Earth interactions.



c) Develop a Strategic Plan.

NASA recently released a new Strategic Plan in 2006 (*2006 NASA Strategic Plan NP-2006-02-423-HQ*). Strategic Goal 3 of this plan addresses Earth science. Within this goal, Sub-goals 3A and 3B are applicable to the GCMD:

Strategic Goal 3: Develop a balanced overall program of science, exploration, and aeronautics consistent with the redirection of the human spaceflight program to focus on exploration.

Sub-goal 3A: Study Earth from space to advance scientific understanding and meet societal needs.

Sub-goal 3B: Understand the Sun and its effects on Earth and the solar system.

Develop Key Aspects of GCMD's Strategic Plan:

Vision

The Global Change Master Directory project envisions becoming the world's premiere source to discover scientific data and provide direct access to selected data sets. Through the wise design and maintenance of internal applications, the GCMD also envisions enabling the discovery of and direct access to relevant data services for the analysis and visualization of data. Project activities are designed to promote the stewardship of scientific resources and thereby build the knowledge and resource connections to enhance our understanding of the Earth, the Solar System, and the Universe.

Mission

The mission of the Global Change Master Directory (GCMD) is to enable the scientific community to discover and access Earth science data and services through distributed, integrated information technology systems. The GCMD offers authoring tools to assist in achieving this mission. These tools are available to write, directly submit, and directly update metadata records, which conform to international standards.

Primary Goals

- *Improve the depth and breadth of scientific research by offering a source of information about data and services.*
- *Increase and maintain the content of the GCMD's unique source for data and services.*
- *Provide direct access to data and services wherever possible.*
- *Continuously improve the search and retrieval of data set and services information through controlled keyword hierarchies, free-text, query refinement, and web site navigation.*
- *Maintain and improve software systems, using state-of-the-art information technologies for the management and delivery of scientific information.*

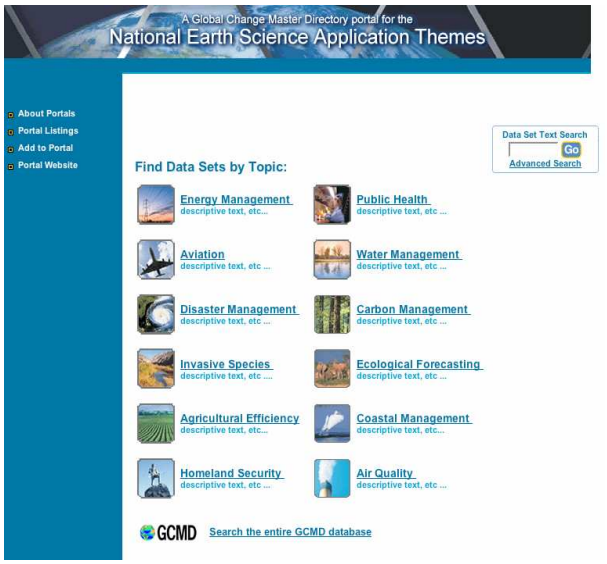
2. Align with NASA ESE science and application themes.

The Group on Earth Observations (GEO), of which the United States Interagency Working Group on Earth Observations (IWGEO) is a component, has developed the 10-Year Implementation Plan for the Global Earth Observation System of Systems (GEOSS).

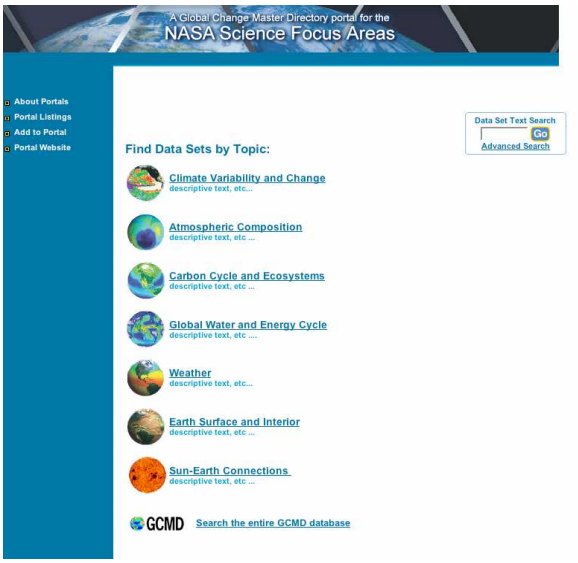
The staff has completed the mapping of the 12 National Application themes and the 6 NASA focus areas to the GCMD Science and Services keywords (where feasible). Federal agency (including NASA) DIFs and SERFs will be automatically populated with one or more of the National Application themes and science focus area so that users will

be able to locate data and services associated with those themes. Mock-up portals have been created based on those themes.

Application Theme portal.



NASA Science Focus Area portal.



a) Re-align content with NASA's priorities.

We are monitoring changes in guidance from NASA HQ to constantly reassess the scope of our content. If we are called upon to add lunar, planetary, astronomy metadata records, our software is flexible and extensible enough to accommodate such content. If there should be a change in the mission of GCMD, then plan to seek expert guidance on astronomical keywords from the National Space Science Data Center (NSSDC), other NASA astronomy divisions, and universities. We have examined astronomical keywords that were originally part of NASA's Master Directory and keywords from the American Geophysical Union (AGU) for potential integration into the GCMD. More expertise in astronomy/planetary science may be needed to establish a set of hierarchical keywords for search and retrieval.

Proposed sample of astronomical keyword hierarchies:

Space Science	Astronomy	Astronomical Parameters	Abundances
Space Science	Astronomy	Astronomical Parameters	Colors
Space Science	Astronomy	Astronomical Parameters	Cross Identifications
Space Science	Astronomy	Astronomical Parameters	Distance
Space Science	Astronomy	Astronomical Parameters	Magnetic Fields

Space Science	Astronomy	Astronomical Parameters	Magnitudes
Space Science	Astronomy	Astronomical Parameters	Masses
Space Science	Astronomy	Astronomical Parameters	Morphology
Space Science	Astronomy	Astronomical Parameters	Object Counts
Space Science	Astronomy	Astronomical Parameters	Occultations
Space Science	Astronomy	Astronomical Parameters	Parallaxes
Space Science	Astronomy	Astronomical Parameters	Photometry
Space Science	Astronomy	Astronomical Parameters	Polarization
Space Science	Astronomy	Astronomical Parameters	Positions
Space Science	Astronomy	Astronomical Parameters	Proper Motions
Space Science	Astronomy	Astronomical Parameters	Radial Velocities
Space Science	Astronomy	Astronomical Parameters	Reddening
Space Science	Astronomy	Astronomical Parameters	Redshifts
Space Science	Astronomy	Astronomical Parameters	Rotational Velocities
Space Science	Astronomy	Astronomical Parameters	Size
Space Science	Astronomy	Astronomical Parameters	Space Velocities
Space Science	Astronomy	Astronomical Parameters	Spectra
Space Science	Astronomy	Astronomical Parameters	Spectrophotometry
Space Science	Astronomy	Astronomical Parameters	Temperature
Space Science	Astronomy	Astronomical Parameters	Variability
Space Science	Astronomy	Astronomical Parameters	

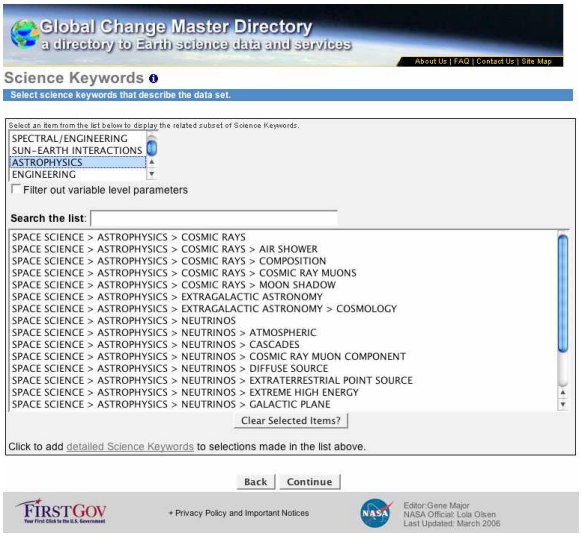
Space Science	Astronomy	Interstellar Objects	Galaxies
Space Science	Astronomy	Interstellar Objects	Nebulae
Space Science	Astronomy	Stellar Objects	Binaries
Space Science	Astronomy	Stellar Objects	Black Holes
Space Science	Astronomy	Stellar Objects	Dwarfs
Space Science	Astronomy	Stellar Objects	Giant
Space Science	Astronomy	Stellar Objects	Interstellar Medium
Space Science	Astronomy	Stellar Objects	Large-scale Structures
Space Science	Astronomy	Stellar Objects	Main Sequence
Space Science	Astronomy	Stellar Objects	Molecular Clouds
Space Science	Astronomy	Stellar Objects	Protostars
Space Science	Astronomy	Stellar Objects	Super Novae
Space Science	Astronomy	Stellar Objects	

Space Science	Planetary Science	Atmospheres
Space Science	Planetary Science	Fields and Particles
Space Science	Planetary Science	Geosciences
Space Science	Planetary Science	Rings
Space Science	Planetary Science	Small Bodies

3. Connect to the space science community.

a) Develop collections that bridge the needs of the space science community.

Sun-Earth science keywords have been revised and new Sun-Earth keywords have been added to reflect data collected from NASA's solar/space missions (e.g., ACE, SOHO, GEOTAIL, WIND, etc.). Two new Sun-Earth Interactions TERMs have been added: Ionosphere/Magnetosphere Dynamics and Solar Energetic Particle Flux. Space/astronomy keywords have been added (see below) by expanding the Science Keyword Category to include SPACE SCIENCE keywords. These additional keywords address the needs of some space science Antarctic researchers and are available through the docBUILDER tools, but are not yet available for searching (see below).



b) Consider opportunities to work with space science community.

Originally, the DIF was designed to accommodate space science data set descriptions. The DIF is remarkably flexible as a scientific discovery metadata format, and the software infrastructure can be adapted to search and discover space science data sets and services.

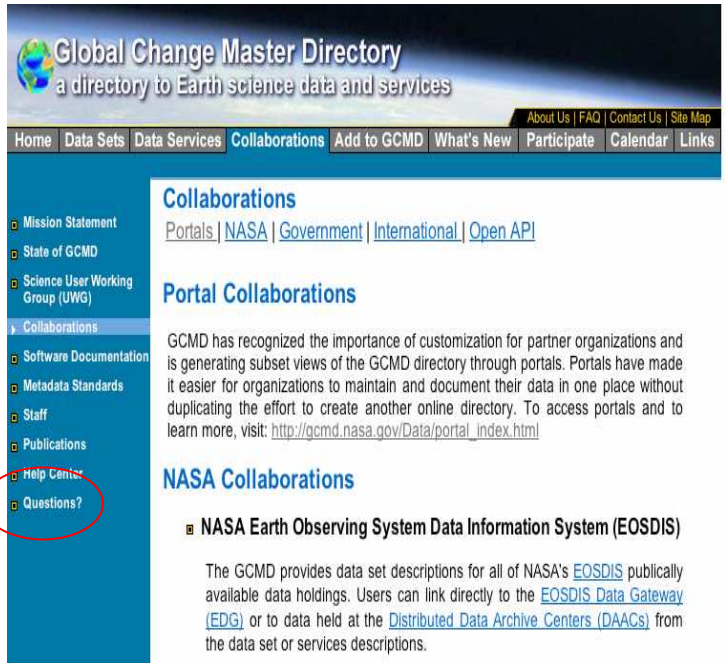
4. Who are the users?

a) Interact more directly with primary users and use interactions to inform decisions on user interfaces and system changes.

Every web page now has an option that allows users to link to the User Comment form to ask questions or provide feedback. Staff members have conducted usability studies and

solicited independent feedback. Characteristics of the MD9.0-9.4 user interface were influenced by the feedback from usability studies and user statistics. See Section 4c.

Link to comment form appears on every web page.



b) Get direct science feedback on usefulness of the directory.

Users are encouraged to provide feedback to the GCMD (see 4a). The form below is used to provide feedback to the GCMD. In 2005, 353 user support questions were received and answered.

c) Invigorate efforts to capture "real" user comments.

The following comments were received through usability tests recommended by the UWG:

- Update navigation tabs
- Improve long scrolling of words
- Remove "go to" anchors
- Include brief summary under titles
- Make free-text query easier to locate in results set

The following changes were made in response to specific user feedback:

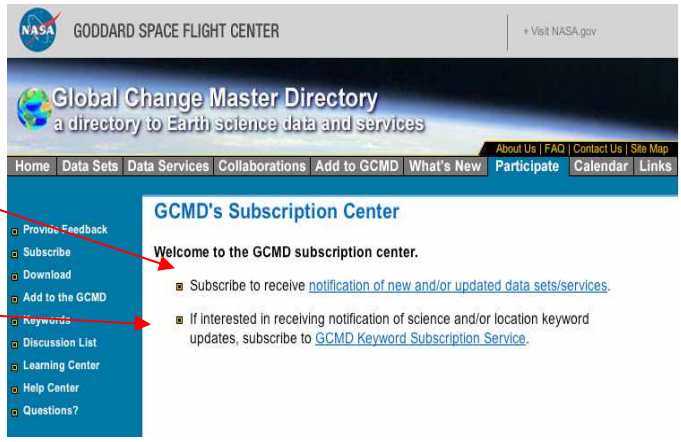
- Modified tab names to *Collaborations* (formerly "Portals"), *Add to GCMD* (formerly "Authoring"), and *Participate* (formerly "Community").
- Made black navigational tabs easier to notice by updating the GCMD header image
- Added detailed information to navigation menu tabs that appear when the computer mouse is hovered over the menu tab name.
- Removed "Go to" anchor links at the top of page.
- Redesigned keyword search pages and added "breadcrumbs" for easier navigation.
- Included brief summary under title display.
- Highlighted free-text search terms in results.
- Added icons for services.
- Added dynamic search terms to the keyword search interface.

d) Consider voluntary user registration to get feedback and encourage questions.

We developed several alternatives to voluntary user registration (see *Section e* below on NASA's policy). These include (1) a subscription service, which is offered for users to receive updates of new DIFs and SERFs, (2) a subscription service for users to receive notification of new and updated controlled keywords, and (3) a comment form for user feedback.

DIF/SERF notification service

Keyword notification service



According to the GSFC webmaster policy (NPR 2800.1), there are limitations on web surveys (see Section 2.3.2), which restrict NASA from collecting user information: http://nodis3.gsfc.nasa.gov/displayDir.cfm?Internal_ID=N_PG_2800_0001_&page_name=Chapter2.

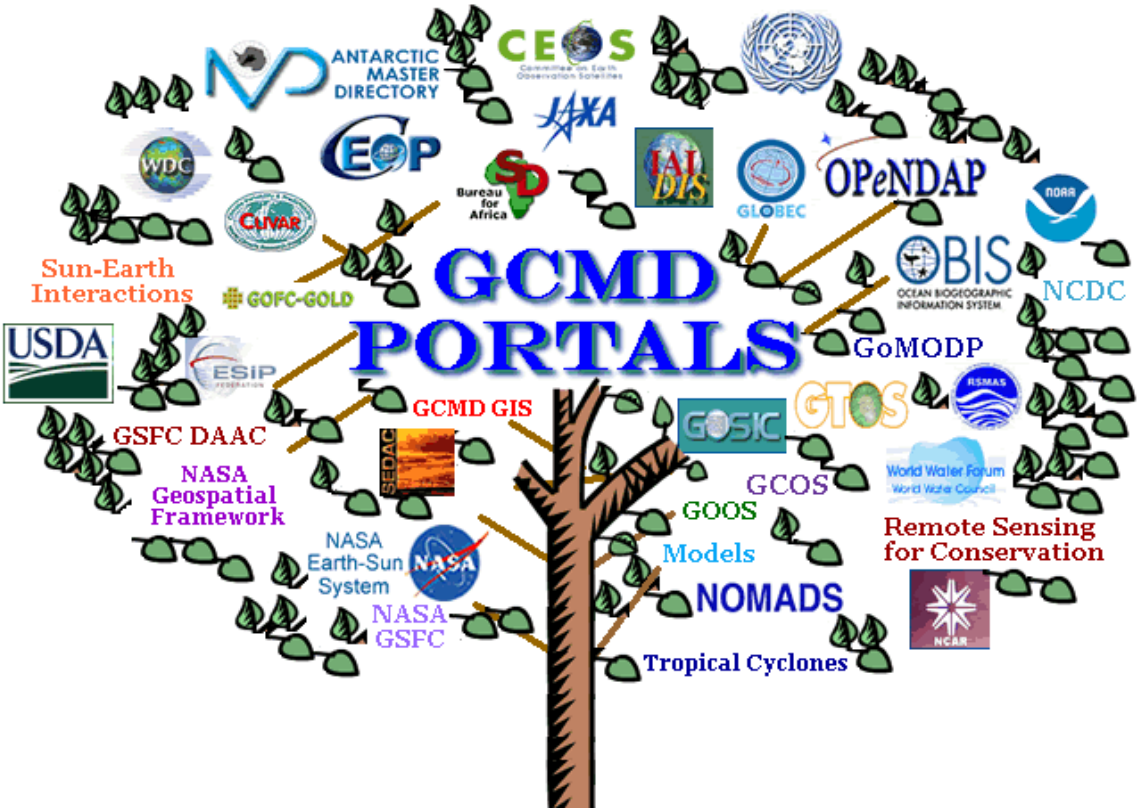
> "2.3.2. Information Collection Clearance Procedures
>
> Any NASA employee proposing collection of information from the public, including from NASA contractors, must justify the requirement, take steps to minimize the collection burden, and comply with clearance and reporting procedures. Before initiating collection of information from the public, the cognizant ICB Clearance Officer must submit a request for clearance to the Agency Clearance Officer."

e) Continue efforts to make connections to professional societies and disciplinary communities.

Posters and papers are presented at professional society meetings such as AGU, AMS, and GSA. We continue to make efforts to reach out to disciplinary communities, such as:

- Gulf of Maine Ocean Data Partnership: requested modifications to the docBUILDER tools such as being able to preview the DIF.
- Ocean Biogeographic Information System: requested accommodation of taxonomic controlled keywords and additional portals for its nodes.
- International Polar Year: interested in the DIF as its metadata standard.
- National Antarctic Data Centers: requested additional portals.
- Task Leader serves on the Executive Board of the Atmospheric Science Librarians International (ASLI).
- Staff are members of the American Geophysical Union, the American Meteorological Society, the Association of American Geographers, the American Society of Limnology and Oceanography, The Oceanography Society.

f) Make portals available to specific communities



We received requests from the science community for new portals. The number of portals has increased from 31 to 67 in the last two years. Since January 2006, the following portals have been added:

- Gulf of Maine Ocean Data Partnership (GoMODP)

Joint Committee on Antarctic Data Management (JCADM) portals:

- AMD_FR
- Arctic_FR

- Committee on Earth Observation Satellites (CEOS) portals:
  - Global Earth Observation System of Systems (GEOSS)
  - Coordinated Enhanced Observing Period (CEOP)

- Ocean Biogeographic Information System (OBIS) portals:

- Australia	- Japan
- Canada	- New Zealand
- China	- South America
- Europe	- Sub-Saharan Africa
- Indian Ocean	- United States

5. Partnerships with science community

a) Inform developing programs of the benefits of GCMD.

The staff met with the NASA/GSFC Software Integration and Visualization Office (SIVO) group to inform them of GCMD's capabilities. The GCMD could assist users in locating modeling data output and gain access to their holdings when they become publicly available. The Ocean Biogeographic Information System (OBIS) community has been very supportive of the directory and is using the GCMD's capabilities for managing and enriching OBIS discovery metadata. They have made extensive use of the GCMD metadata tools and have provided valuable user feedback. Representatives from OBIS are demonstrating the functionality and benefits of the GCMD to outside partners. The GCMD is also collaborating with the Gulf of Maine Ocean Data Partnership (GoMODP) to promote and coordinate the sharing, linking, and use of data in the Gulf of Maine region. The GCMD has been participating in the International Polar Year (IPY) data and information activities. We have informed the IPY of the benefits of using an established directory and metadata format for discovering polar data sets, participated in writing an IPY proposal, and participated in an IPY meeting/workshop in Cambridge, England in March 2006.

b) Work closely with scientists who develop data sharing proposals and have GCMD offer service to create infrastructure to offer the data.

We wrote two NASA ROSES proposals:

- Accessing NASA's Satellite Data for GIS Analysis
- Adaptation of Earth Science Software Ensemble for Planetary Exploration

c) NASA should encourage and/or require all its projects to contribute metadata to GCMD.

NASA has an explicit requirement for all ReASON CAN awardees and original ESIP Federation awardees to provide metadata to the GCMD. In addition, the ESIP Federation has a less stringent policy that requests all members (beyond the original awardees) to provide metadata to the GCMD. Staff members actively participate in ESIP Federation subcommittees. In 2005, the GCMD was accepted as a full ESIP member, thereby solidifying our presence within the Federation.

d) Choose collaborations carefully.

Collaborations are targeted that may yield the greatest benefit in terms of growth and exposure to the mission and goals of the GCMD, without compromising resources and assets and are coordinated with our priorities. The top 10 priorities, upon which we base our collaborations are:

1. Satisfy funding agency obligations first.

2. Become the most comprehensive and accurate data/services database in the world (for Earth science).
3. Provide 24/7 continuous operations with automatic backup support, and iron-clad tight security.
4. Host a flawless web site, with intuitive navigation, avoiding outdated content and misinformation.
5. Provide direct access to data and services, including the handling of geospatial data.
6. Reduce overall maintenance of content, software, and hardware.
7. Release software only after four levels of testing are complete: developer testing, full project testing, external testing with collaborators, and external testing with known outside sources.
8. Continue to grow and evolve technologically to maintain an "edge" on the competition; be cognizant of international standards.
9. Devise clever new functionality. For example, provide near real-time temporal coverage option.
10. Become recognized as a reliable source for superb content and software, as well as for our willingness to share.

Collaboration criteria:

Based on the UWG recommendation to choose collaborations carefully, we have established criteria to assist us in our prioritization. We plan to ask ourselves the following questions to prioritize any collaboration in light of other requests if time and resources are available.

1. Does the collaboration satisfy HQ and other existing funding sources?
2. Does the collaboration follow or fit with NASA's mission?
3. Does the collaboration include a clearly-stated MOU regarding the expected role of the collaborator and offer resources necessary to fulfill this role?
4. Does the collaboration come with funding? If so, how will funds be transferred to the project?
5. Exactly what is expected/required of the GCMD in the collaboration? An outline of the exact requirements expected from the GCMD is needed.

    What are the reputations of the other parties involved in the collaboration and how will those reputations reflect on the GCMD?

    What is required of the requestor of the collaboration? How will the requestor be participating in the collaboration?

6. Will the collaboration allow us to increase our data and services content? If so, what is the estimated cost associated with this?
7. Will NASA/GCMD be properly credited as the originator of the content and/or the software products? Do the collaborators expressly desire to include the GCMD in the final product?

8. Would the collaboration give GCMD added exposure/recognition from a well-established agency or project in the community in terms of others that may then consequently provide some sort of positive repercussion?
   Does the collaboration bring us increased exposure, in terms of garnering new users? Is it worth the expenditures on our part for expected results?
9. Does the collaboration offer an opportunity to extend our expertise in an area previously identified in our future plans? Example: OAI, web services.
10. How do we know when a collaboration is no longer useful?
    Does the collaboration have a short, medium, or long-term goal? If so, what is the limit of the collaboration in terms of time?

e) Focus on NASA-related conferences, as well as metadata and library conferences.

The staff participates in the GSFC Library Metadata Review Group (MRG) by advising the MRG on the creation of the Goddard Core Descriptive Element Set (a library resource metadata format based on Dublin Core). At our suggestion, they have included geospatial elements. We also participate in the Earth Science Information Partnership (ESIP) metadata subcommittees and conferences, as well as in Global Organization for Earth System Science Portal (GO-ESSP), the Metadata Marine Interoperability (MMI), the Integrated Ocean Observing System/Data Management and Communications (IOOS/DMAC), and the Observation Working Group/Observation Working Group Data and Information System (OWG/OWGDIS) of the Climate Change Science Program (CCSP). At least one or more abstracts are submitted every year to major science conferences such as the American Meteorological Society (AMS), American Geophysical Union (AGU) and Geological Society of America (GSA). An invited paper was given at the Atmospheric Science Librarians International (ASLI) meeting held in conjunction with the American Meteorological Society (AMS) in January 2006. Another invited paper was presented at the Computing in Atmospheric Sciences Conference (CAS2K5) in September 2005. When possible, staff participates in demonstrating the GCMD at the OneNASA booth at major conferences.

6. Directory vs. catalog and direct access to data.

Direct access to data and services is provided through the Related_URL field within the DIF and SERF metadata. The Related_URL types were examined and reduced from 26 to 14 categories. Several URL_Content categories have been mapped directly to the new "Get Data" or "Get Service" categories, which inform the user that the URL will link them directly to the data or service described by the DIF or SERF.

Current Keywords	New Keywords (Proposed)
ASSOCIATED DATA SET(S)	GET DATA
BDP METADATA	VIEW EXTENDED METADATA
DISC-AIRS	GET DATA
DISC-MODIS	GET DATA
DISC-OMI	GET DATA
DODS_DIR	GET DATA
DODS_LAS	GET DATA
DODS_URL	GET DATA
EML METADATA	VIEW EXTENDED METADATA
FGDC METADATA	VIEW EXTENDED METADATA
GET DATA	GET DATA
GET SERVICE	GET SERVICE
LAS	GET DATA
NBII METADATA	VIEW EXTENDED METADATA
NOMADS DATA SERVER	GET DATA
PROJECT HOME PAGE	VIEW PROJECT HOME PAGE
REFERENCE MATERIALS	VIEW RELATED INFORMATION
RELATED DATA SET DESCRIPTION	VIEW RELATED INFORMATION
RELATED DATA SET METADATA	GET RELATED DATA SET METADATA (DIF)
RELATED SERVICE METADATA	GET RELATED SERVICE METADATA (SERF)
SOAP API	ACCESS WEB SERVICE
SOFTWARE PACKAGES	GET SOFTWARE PACKAGE
UDDI REGISTRY	ACCESS WEB SERVICE
WEB COVERAGE SERVICE (WCS)	GET WEB COVERAGE SERVICE (WCS)
WEB FEATURE SERVICE (WFS)	GET WEB FEATURE SERVICE (WFS)
WEB MAPPING SERVICE (WMS)	GET WEB MAPPING SERVICE (WMS)
NEW	GET MAP SERVICE
NEW	ACCESS MAP VIEWER

7. Keyword brokering

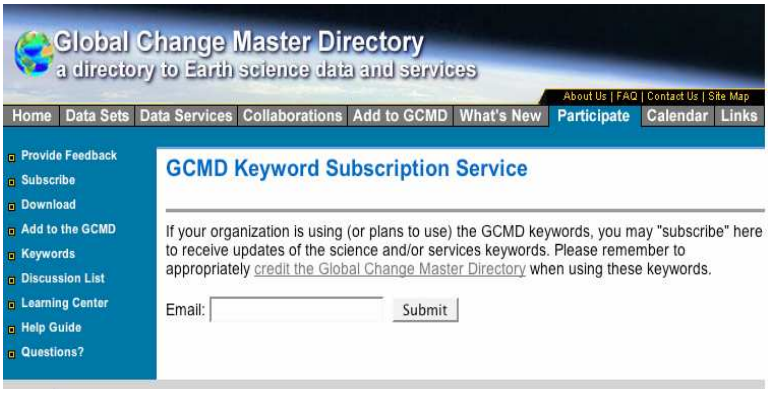
a) Establish role as keyword and controlled vocabulary broker.

By offering an official citation, keyword rules, and procedures for the modification of keywords, we are finding greater acceptance.

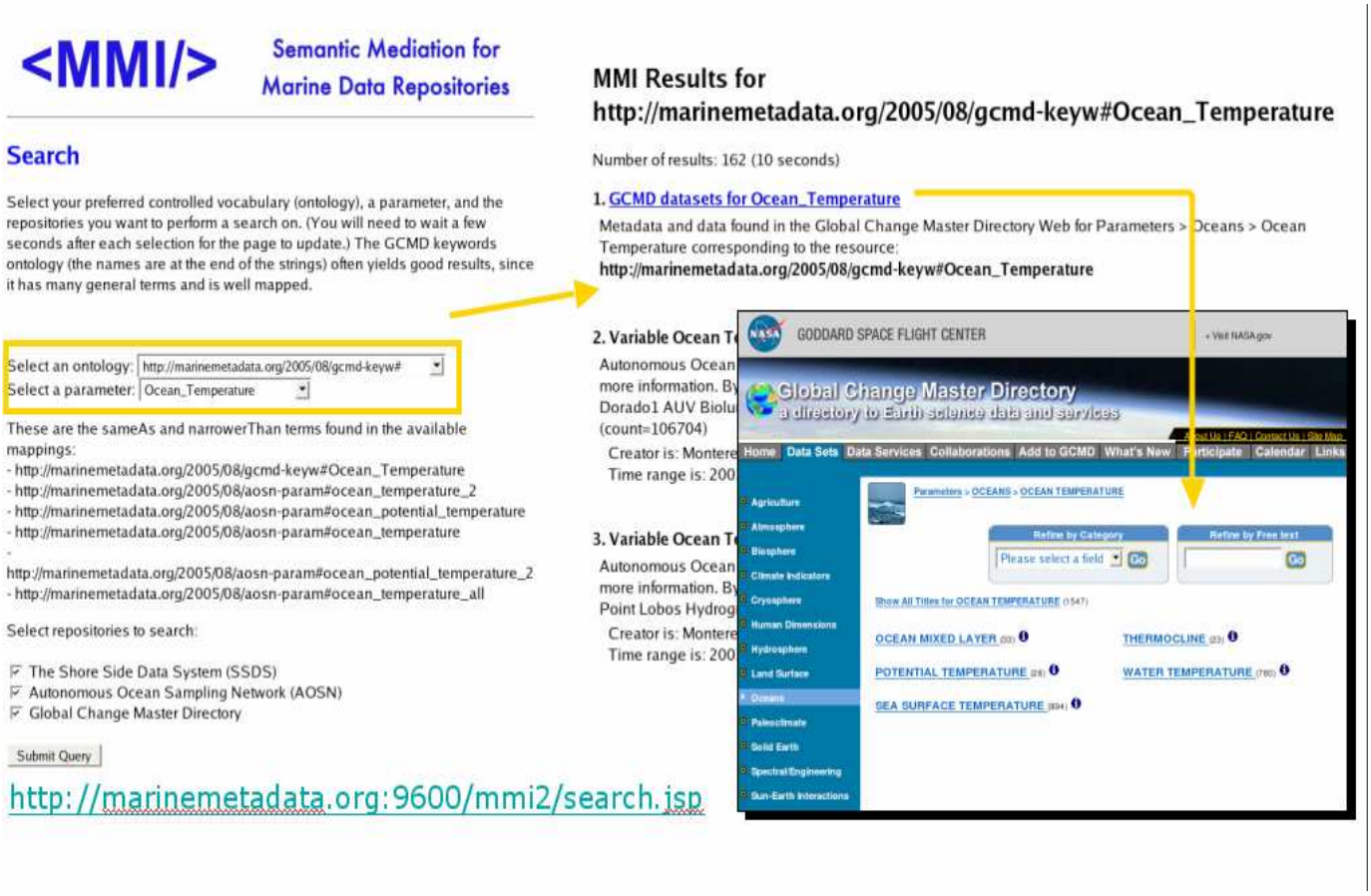
b) Encourage use of keywords. Take the high road when users ask for keywords.

We have implemented a keyword subscription service (see also Section 4d) that allows the user to be notified (via email) when new science keywords have been added or modified. Detailed keywords have also been updated for consistency and spelling errors. By offering a keyword subscription service, we encourage use of the GCMD controlled

keywords. This email service helps users keep informed of new and updated controlled keywords and also encourages users to synchronize their keywords with the latest GCMD keywords. See the State of the GCMD 2005 report for a complete list of these known organizations that are using the GCMD keywords.



Example: Use of GCMD keywords with the Marine Metadata Initiative.



c) Make explicit reference to citation for keywords so that users can give proper NASA attribution.

Citation information for our keywords already exists on the GCMD web site (http://gcmd.nasa.gov/Resources/valids/)

> Olsen, L.M., G. Major, S. Leicester, K. Shein, J. Scialdone, H. Weir, S. Ritz, C. Solomon, M. Holland, R. Bilodeau, T. Northcutt, R. Vogel. 2006. *NASA/Global Change Master Directory (GCMD) Earth Science Keywords*. Version 5.3.3, online: http://gcmd.nasa.gov/Resources/valids/keyword_list.html]

d) Consider extending keywords to include space science topics

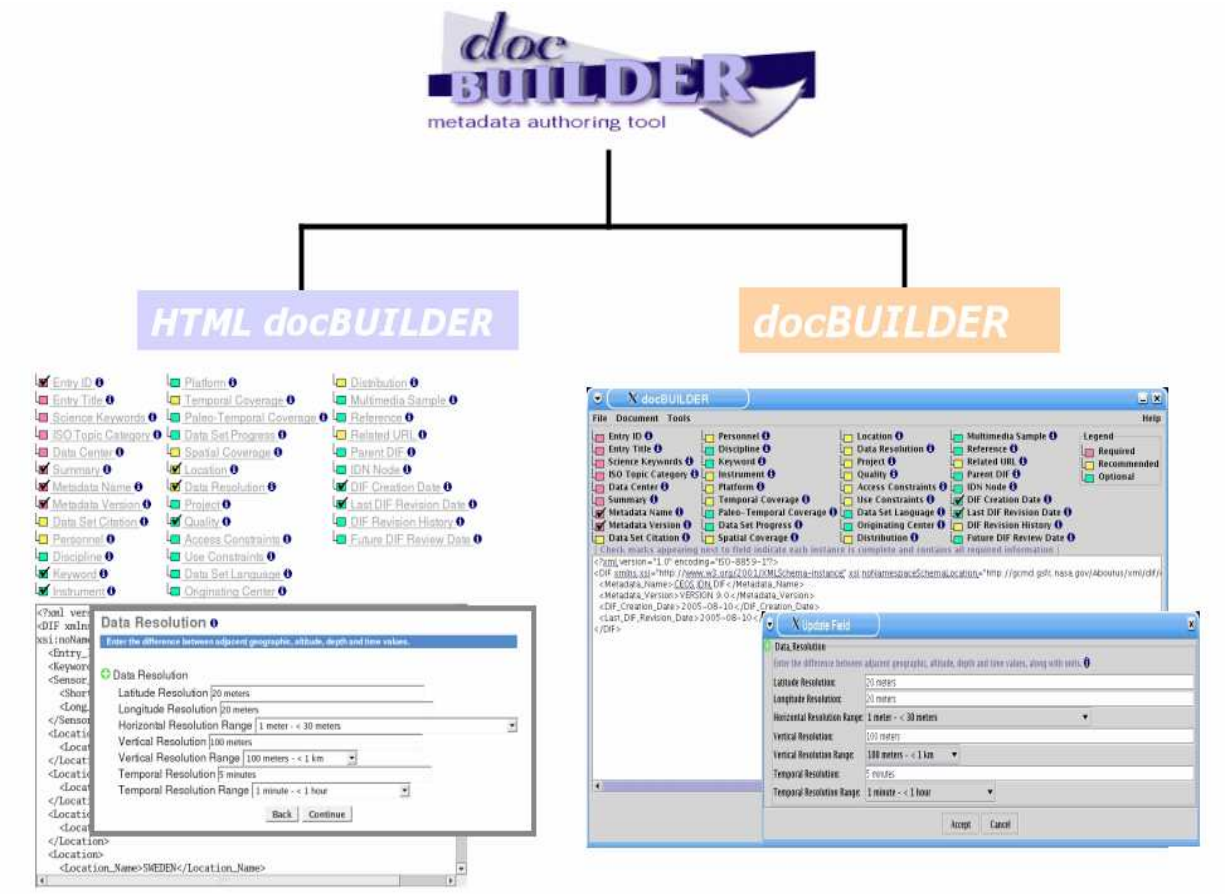
We have added a "Space Science" Category and have populated a few keywords to accommodate the Antarctic astrophysics community. More work and expertise is needed to create space keyword hierarchies. (See Section 3 a).

8. Community contribution and maintenance of metadata

a) Encourage data providers to contribute and maintain their metadata, including updates.

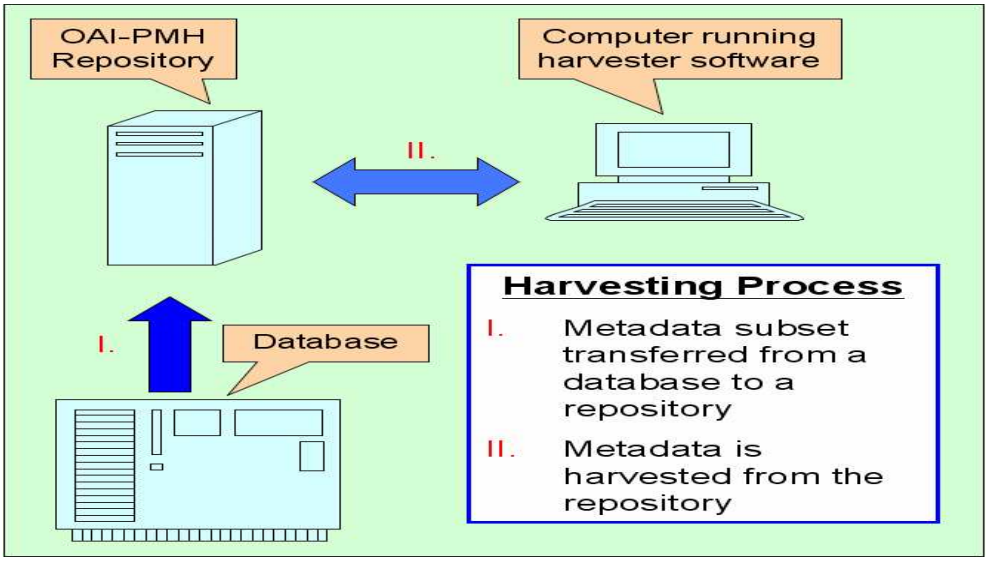
The docBUILDER tools have been upgraded and a new stand-alone docBUILDERsolo tool has been implemented for users to easily create DIF and SERF metadata records. It should be noted that many collaborators regularly use the docBUILDER tools: Antarctic researchers, DAACs, ESIP members, NOAA, and GLOBEC. The docBUILDER highlights include:

- Refactored framework using Struts and JSP allowing for complete portal customizations and a clean separation of business and presentation logic.
- Refactored business logic with a completely rewritten business layer resulting in much faster access to XML documents and valids.
- Support for templates.
- Improved Data Center and Personnel HTML form including a more intuitive design for selecting Data Centers and Personnel.
- Save Edits Feature, allowing users to edit XML documents directly through the tool.
- Downloadable offline version of docBUILDER tool (docBUILDERsolo)
- Improved JavaScript widget for selecting valids.
- 3 color support (required, optional, not required).
- Automatic population of Data Center URLs.
- Spell Check feature available in offline version.
- Compatible with XML Schema.

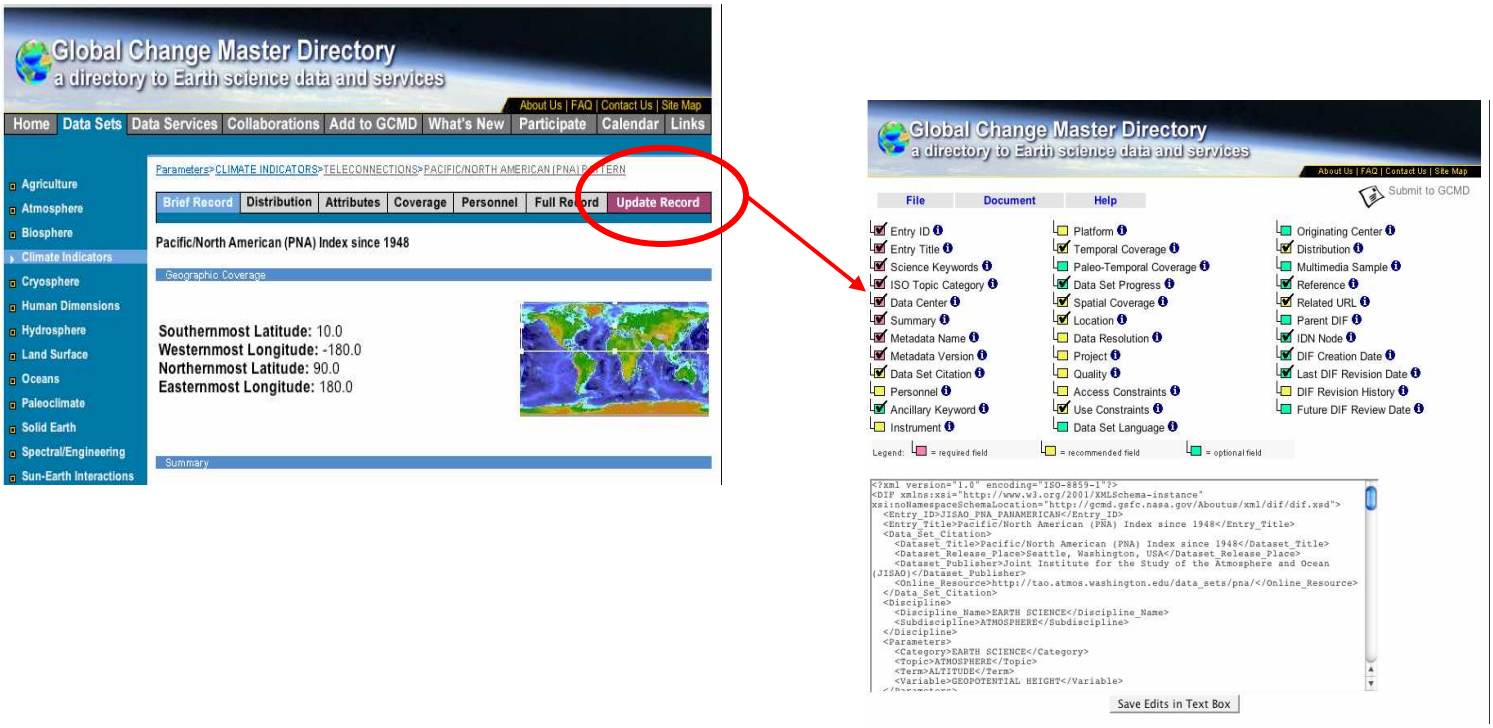


In addition to the docBUILDER tools, the staff has also taken advantage of web technologies such as eXtensible Stylesheet Tranformation (XSLT) that make it easier to translate and import large quantities of metadata in formats other than DIF. We have been working closely with NOAA/NCDC to successfully harvest metadata (in FGDC format) from their NOAA Metadata Manager Repository (MMR) system using XSLT, and we regularly translate metadata from GeoConnections (Canada).

We have also experimented with harvesting technologies using the Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH) as another way for partners to contribute metadata. During the summer of 2005 test metadata harvests were conducted between the GCMD and NCAR Community Data Portal (CDP) OAI-PMH servers and in early 2006, metadata was harvested from THREDDS servers.



Updates to DIFs and SERFs can be made through the “Update” tab on the DIF or SERF display making it easy for providers to maintain their metadata.



b) Explore ways to teach the community how to develop their own DIFs and SERFs.

Through conferences and presentations, the GCMD staff has demonstrated to data and service providers how to write their own metadata records. A web capture video demonstration has been prepared (.avi file) to illustrate how to create a metadata record. Presentations at AGU, ESIP and CEOS meetings demonstrated how docBUILDERsolo, the new standalone version of docBUILDER, could be used when not connected to the Internet. This tool should prove to be very beneficial to the members of the Antarctic community for writing metadata while they are in Antarctica, allowing them to save the metadata locally and later resubmit it to the directory.

c) Implement automated reminders.

A functionality to automatically remind authors to update their metadata will be released with MD9.6. A sample email is displayed below:

TITLE: GCMD Metadata Author Reminder

Dear Metadata Author,

This email has been sent to inform you that your metadata records have not been updated in two years and may need to be updated. We ask that you review the following records to ensure the information supplied is still current.

In addition, we ask that you consider:

A. Adding a direct link to your data or service through the Related_URL filed or verifying a previously supplied link.
B. Populating the new field now available for data set resolution.
C. Adding or verifying the geographic coverage of your data through the Spatial Coverage field (A new and improved search for this coverage is now available.)

(To update a record follow the links below and click on the purple "Update Record" tab.) If you require assistance, please do so through http://gcmd.nasa.gov/MailComments/index.html

Offshore, Coastal and Selected Lakes Geophysical Survey Tracks - Expedition Database [CANEMRCCRSMULTIPARAM]
http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&EntryId=CANEMRCCRSMULTIPARAM&MetadataView=Brief&MetadataType=0

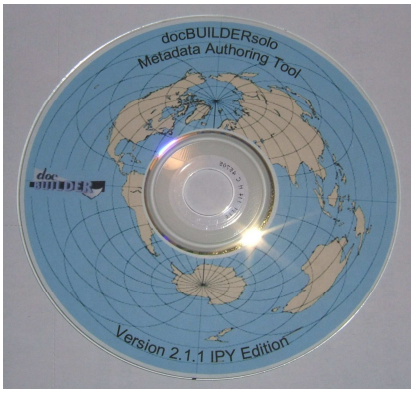
ICEMELT LINE [PASSCAL_ICEMELT]
http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&&EntryId=PASSCAL_ICEMELT&MetadataView=Brief&MetadataType=0

Model-based assimilation methods for spatially-resolved zooplankton data [GLOBEC_057_UK_007]
http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&EntryId=GLOBEC_057_UK_007&MetadataView=Brief&MetadataType=0

d) Advertise docBUILDER more broadly.

We have advertised docBUILDER tools at conferences and custom docBUILDER CDs were created for distribution at:

- Joint Committee on Antarctic Data Management (JCADM) IX, July 2005
- Computing in Atmospheric Sciences (CAS2K5) Conference, September 2005
- International Polar Year (IPY) Workshop, March 2006



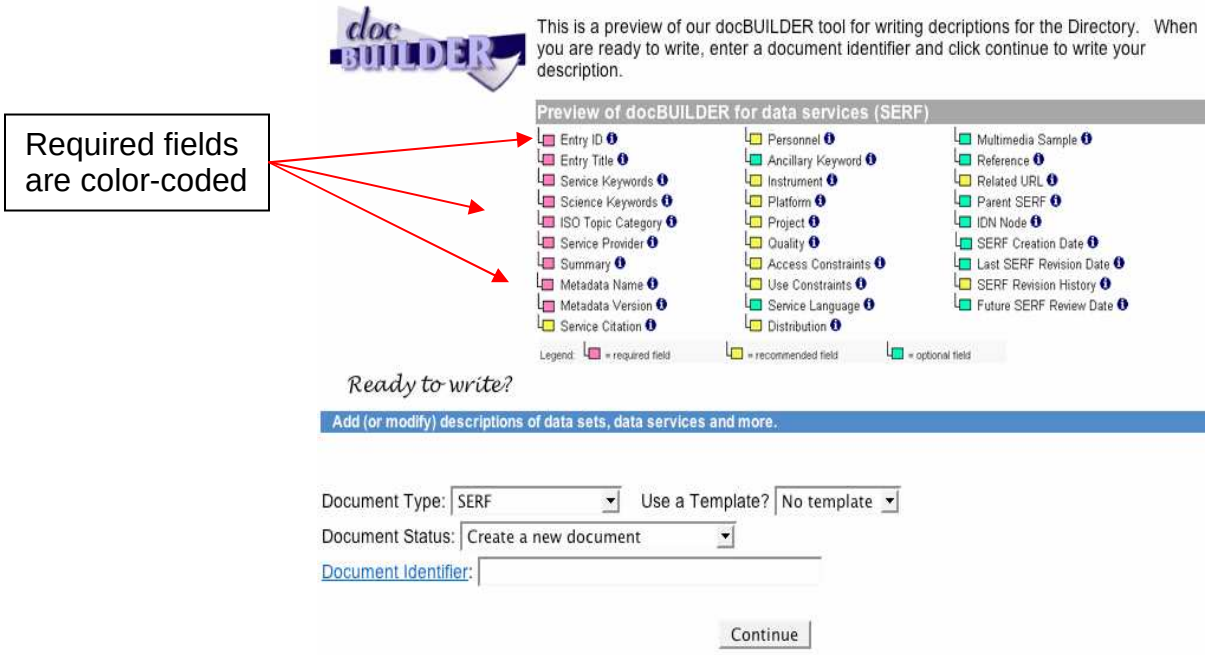
e) Pursue collaborations with partners committed to providing and updating metadata.

In a recent collaboration with NOAA/NODC, we obtained AVHRR Pathfinder SST FGDC metadata. A customized XSLT stylesheet was created for XML translation. This small set of metadata (102 documents) serves as a testbed for further GCMD-NODC collaborations. Other recent collaborations include the European Directory of Marine Environmental Datasets (EDMED), the Gulf of Maine Ocean Data Partnership (GoMODP), and the Ocean Biogeographic Information System (OBIS). A small set of EDMED ISO 19115 metadata was provided to the GCMD by the French Polar Institute (IPEV) and included into the Antarctic Master Directory (AMD) in the summer of 2005. A customized EDMED ISO-to-DIF XSLT stylesheet was created. Additional EDMED

southern ocean entries in the AMD are expected in the future. GoMODP partners are also committed to provide and update metadata. The partnership's metadata contribution increased from five (October 2005) to 114 (December 2005) in just 3 months. At the end of 2005, OBIS finalized plans to improve its discovery metadata and proposed to utilize the GCMD authoring tools and portals. Over 200+ OBIS records are expected in 2006. Dedicated long-time partners such as GeoConnections (Canada), NOAA NMMR, the Joint Committee on Antarctic Data Management (JCADM), the Global Ocean Ecosystems Dynamics Project (GLOBEC), the Global Observing Systems Information Center (GOSIC) and the Federation of Earth Science Information Partner (ESIP) continue to be actively involved in providing and updating metadata in the GCMD. Last year, the AMD portal reported a 25% growth in content from January to December 2005. The growth can be attributed to the continued support and involvement of existing and new National Antarctic Data Centers (NADCs). The GCMD continues to work with those partners and the DAACs to keep the most up-to-date information online.

f) Consider developing a SERF-lite.

We believe our metadata authoring tools automatically offer a “lite” feature. For services, there are *nine* required fields that must be completed to create a SERF. Of those nine, *three* required fields are automatically completed by docBUILDER. The remaining *six* required fields provide valuable information to the user about the tool, resource, or service. The SERF cannot be much "lite-r" than this. With the required information, recommended fields can be completed during the SERF loading process.



9. Semantic web and common metadata standards

a) Pursue policies to encourage common metadata standards beyond NASA.

The eXtensible Stylesheet Language Transformations (XSLT) have been coded to make the DIF compatible with all major metadata standards including FGDC (and its variations), Dublin Core, ISO-19115/TC-211 and ISO-19139 (XML schema for ISO-19115).
The GCMD is actively engaged in exploring metadata standards for interoperability:

- OAI-PMH for harvesting metadata. This has been used with NCAR, NOAA, THREDDS, and ESIP Federation
- W3C standards: XML and XSLT. XSLT is routinely used in transferring metadata from GeoConnections (Canada), NOAA/NMMR, Geospatial One Stop, JCADM (AMD/FR), OBIS, NOAA/NODC, EDMED, GoMODP
- FGDC Remote Sensing extension from NGDC
- FGDC NBII/Biological Profile from OBIS

b) Advance cataloging and metadata collaborations from NOAA and NSF cyberinfrastructure programs such as NMMR, THREDDS, NSDL, DLESE.

In January 2006, the entire collection of NCDC FGDC XML documents stored in the NOAA NMMR (NOAA Metadata Managers Repository) was transferred to the GCMD, translated to DIF XML format, and entered into the database. There are currently 484 NCDC DIFs in the GCMD. The protocol used for harvesting NMMR metadata is currently Unix Rsync. We are hopeful that NOAA will adapt the Open Archives Inititative Protocol for Metadata Harvesting (OAI-PMH) as a simple, standardized, and open-source protocol for metadata harvesting (see Section 8 a). We have tested harvesting NCAR metadata using the Digital Library for Earth Science Education (DLESE) OAI-PMH metadata harvester software. Other potential OAI-PMH partners in 2006 are Geospatial One Stop, NSIDC, and the NASA Langley Research Center.

Recently, THREDDS announced the release of their new catalog system, which will contain "enough metadata to (at long last) to write DIF and ADN digital library records and import those into GCMD and DLESE discovery centers." We have successfully used OAI-PMH to harvest 27 DIFs from THREDDS servers and entered these into the GCMD database. A data discovery meeting, organized by Peter Cornillon, is scheduled for May 2006, which will bring to together many of the data providers such as THREDDS, OPeNDAP, ADL, ECHO, and GCMD.

10. Connecting to DLESE, NSDL and other educational programs

a) Work more closely with NSDL and DLESE communities.

DLESE has an entry for GCMD in their catalog; however, metadata in DLESE is fairly minimal (metadata is based on a modified Dublin Core format). DLESE also utilizes the OAI-PMH protocol, and we have adopted their OAI-PMH provider/harvester software for our experiments with NCAR. It should be noted that DLESE mainly describes educational web sites rather than products.

b) Every SERF could be a DLESE record.

DLESE's focus is on educational web sites and resources for teachers, whereas SERFs cover a much broader area (models, tools, advisories, etc.). We plan to share all educational descriptions with DLESE. See part (a) above.

c) DIFs could reference educational links and SERFs.

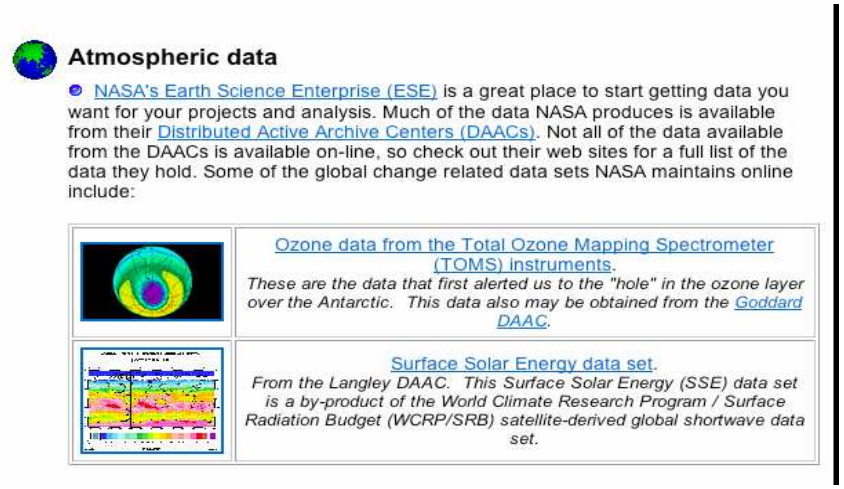
The GCMD currently holds 308 education and outreach services descriptions, and 86 educational resources from the "links" page. Links exist between DIF to SERF (and SERF to DIF) metadata through the Related_URL field. Relevant URL_Content_Type(s) include:

GET RELATED DATA SET (DIF) METADATA
GET RELATED SERVICE (SERF) METADATA

d) Organize educational content.

The GCMD has a "Learning Center" on the web site, which points to NASA data sets that can be readily used as well as links to educational resources.

An example of a link to data you can use in the classroom.



The GCMD has three levels of education/outreach keywords to organize educational descriptions and aid the user in discovering education and outreach services.

Education/outreach > Curriculum Support > Background Information
Education/outreach > Curriculum Support > Classroom Activities
Education/outreach > Curriculum Support > Lesson Plans
Education/outreach > Exhibit Materials > Museum Exhibits
Education/outreach > Exhibit Materials > Science Center Exhibits
Education/outreach > Interactive Programs > Stand Alone
Education/outreach > Interactive Programs > Web-based

e) Consider adding an expert level ranking scheme.

We do not believe we are qualified to rank content (suggested ranking from 9=specialist to 1=Kindergarten). We do not have the resources to undertake this effort and believe the data provider would be more qualified to provide such ranking.

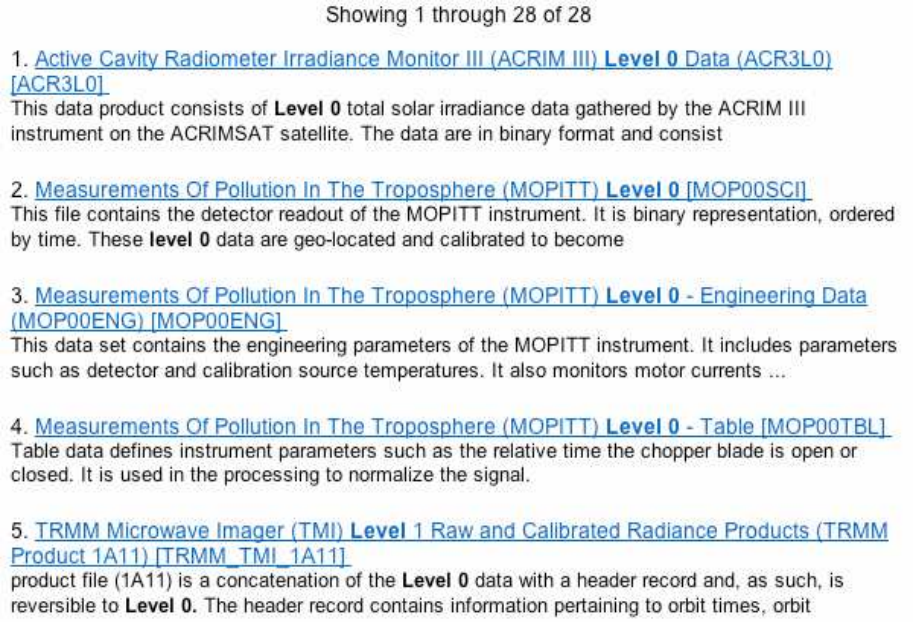
11. NASA data first, and current data first

a) Focus on NASA data first, especially level 0 and 1 satellite data.

Our focus is to obtain data set content descriptions from the NASA DAACs, and other NASA centers (including the ESIP Federation), which include Level 0 and 1 data set descriptions where available. There are 113 NASA Level 1 data sets and 28 NASA

Level 0 data sets. Almost 25% of the GCMD holdings are from NASA centers and partners.

Sample Level 0 NASA data sets.



Showing 1 through 28 of 28

1. Active Cavity Radiometer Irradiance Monitor III (ACRIM III) **Level 0** Data (ACR3L0) [ACR3L0]
This data product consists of **Level 0** total solar irradiance data gathered by the ACRIM III instrument on the ACRIMSAT satellite. The data are in binary format and consist

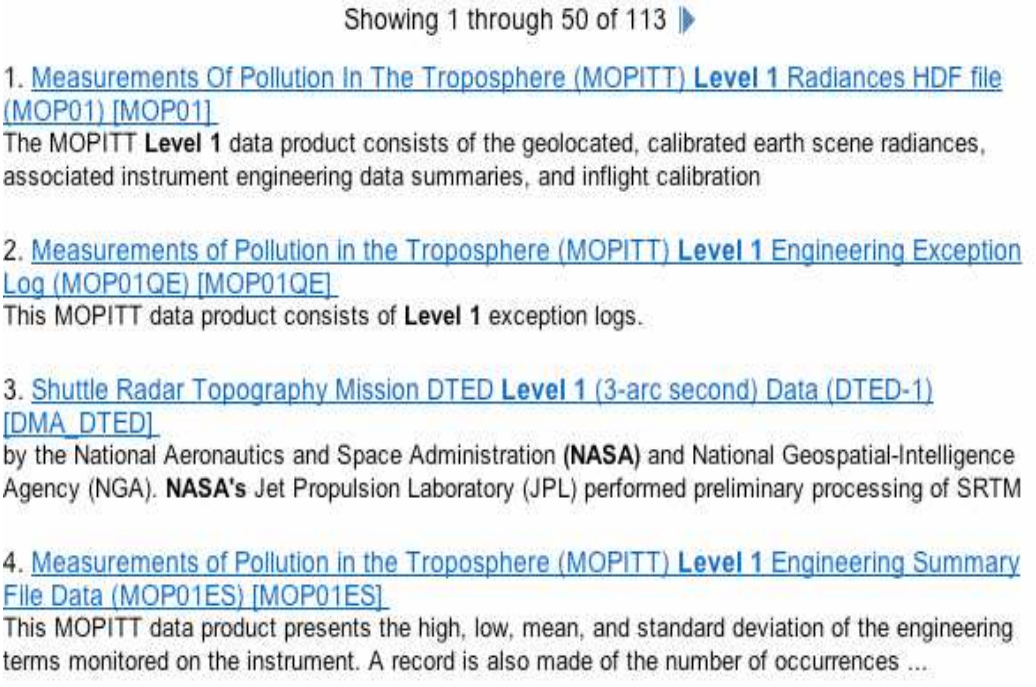
2. Measurements Of Pollution In The Troposphere (MOPITT) **Level 0** [MOP00SCI]
This file contains the detector readout of the MOPITT instrument. It is binary representation, ordered by time. These **level 0** data are geo-located and calibrated to become

3. Measurements Of Pollution In The Troposphere (MOPITT) **Level 0** - Engineering Data (MOP00ENG) [MOP00ENG]
This data set contains the engineering parameters of the MOPITT instrument. It includes parameters such as detector and calibration source temperatures. It also monitors motor currents ...

4. Measurements Of Pollution In The Troposphere (MOPITT) **Level 0** - Table [MOP00TBL]
Table data defines instrument parameters such as the relative time the chopper blade is open or closed. It is used in the processing to normalize the signal.

5. TRMM Microwave Imager (TMI) **Level** 1 Raw and Calibrated Radiance Products (TRMM Product 1A11) [TRMM_TMI_1A11]
product file (1A11) is a concatenation of the **Level 0** data with a header record and, as such, is reversible to **Level 0.** The header record contains information pertaining to orbit times, orbit

Sample of Level 1 NASA data sets.



Showing 1 through 50 of 113

1. Measurements Of Pollution In The Troposphere (MOPITT) **Level 1** Radiances HDF file (MOP01) [MOP01]
The MOPITT **Level 1** data product consists of the geolocated, calibrated earth scene radiances, associated instrument engineering data summaries, and inflight calibration

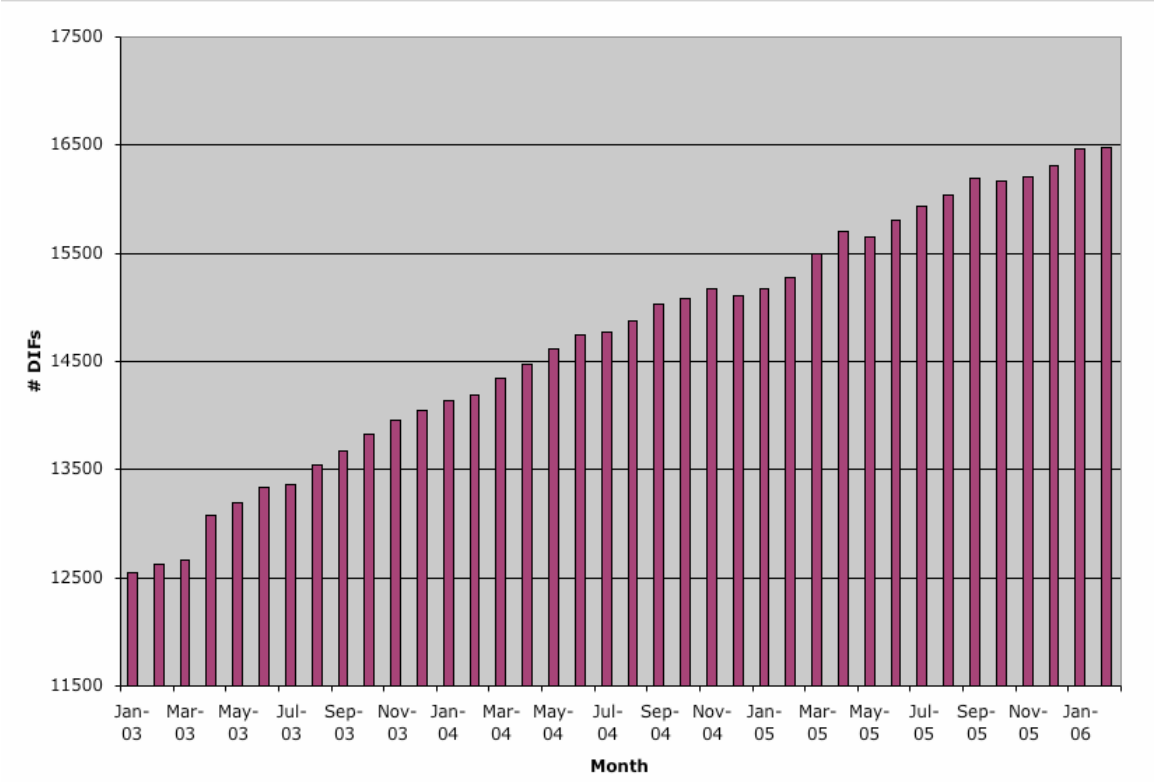
2. Measurements of Pollution in the Troposphere (MOPITT) **Level 1** Engineering Exception Log (MOP01QE) [MOP01QE]
This MOPITT data product consists of **Level 1** exception logs.

3. Shuttle Radar Topography Mission DTED **Level 1** (3-arc second) Data (DTED-1) [DMA_DTED]
by the National Aeronautics and Space Administration **(NASA)** and National Geospatial-Intelligence Agency (NGA). **NASA's** Jet Propulsion Laboratory (JPL) performed preliminary processing of SRTM

4. Measurements of Pollution in the Troposphere (MOPITT) **Level 1** Engineering Summary File Data (MOP01ES) [MOP01ES]
This MOPITT data product presents the high, low, mean, and standard deviation of the engineering terms monitored on the instrument. A record is also made of the number of occurrences ...

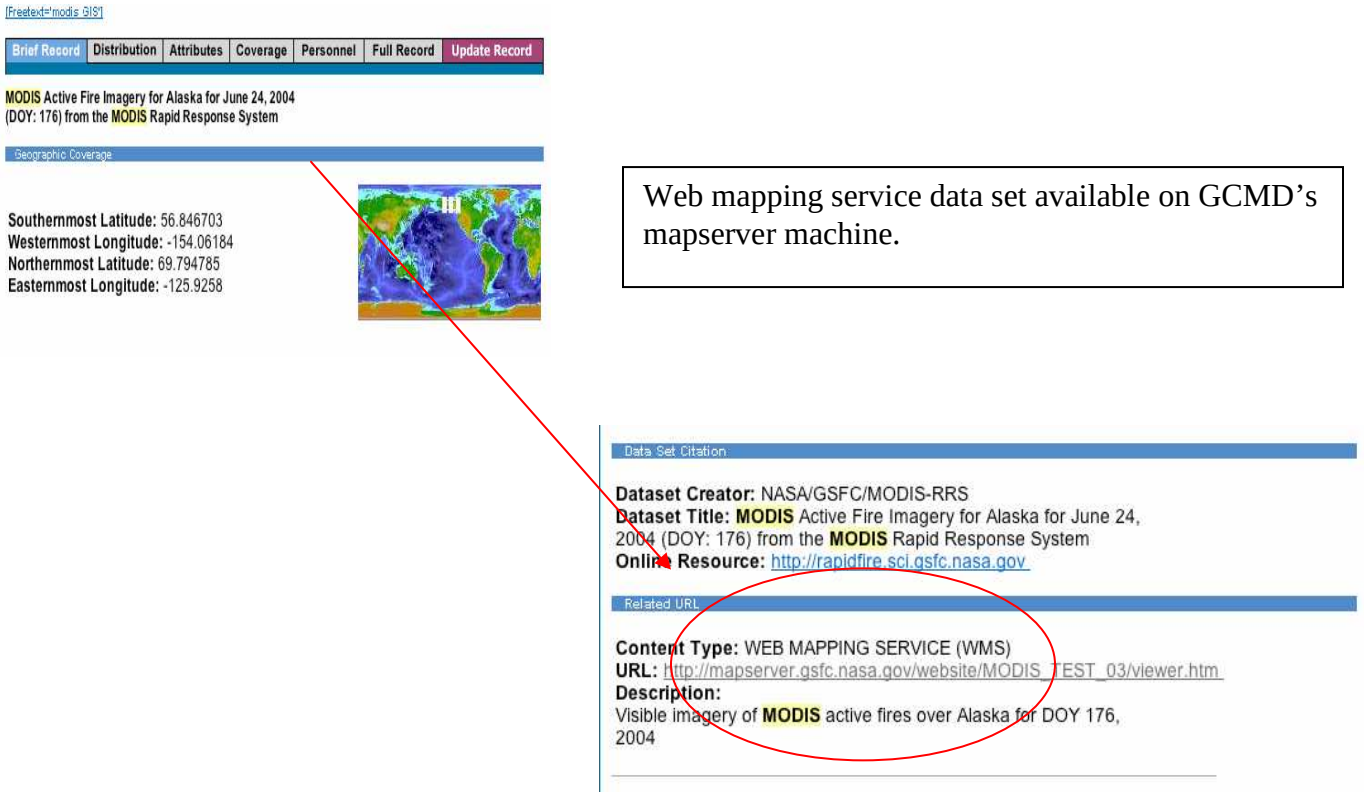
b) Give high priority for new data.

NASA DAACs and ESIP Federation partners are always first priority for new data. The chart below illustrates the cumulative growth in new DIFs since 2003. NASA DAAC and ESIP partners account for nearly 25% of the GCMD database.



12. Explore ways to make small data sets available on the GCMD site.

A computer has been set up (pi-data.gsfc.nasa.gov) at GSFC using Zope/Plone, which hosts a 400G RAID disk for handling small data sets when they become available. In addition, the GCMD has implemented a prototype mapserver, which contains small geospatial data sets provided by the NASA MODIS Rapid Response System, NASA Visible Earth, NOAA, USDA, and ESRI. These data sets can be accessed directly from the DIF.



13. Advisory Board

The UWG has been advised that we continue to use the term UWG in preference to "GCMD Advisory Board".

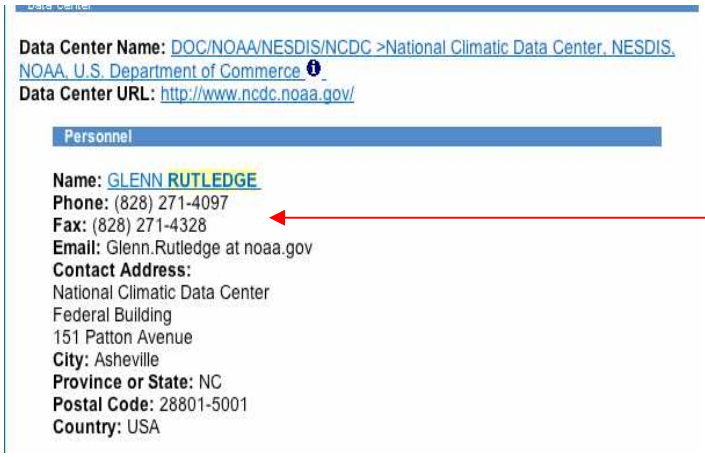
II. Operational Recommendations

1. Robots and search engines: Allow search and indexing robots to access content.

In August 2004, a separate and dynamically generated cache of DIFs and SERFs was made available for Internet robots, such as Google, to index. As a result, usage of the GCMD increased dramatically to an average of over 40,000 unique hosts/month and nearly 5M hits per month. In just one month from August 2004 to September 2004, the number of unique hosts rose 62% and the number of hits rose 122%!

2. Track individuals responsible for data.

Tracking individuals responsible for the data can be a maintenance burden. Email addresses and phone numbers often change. Individuals also change institutions and affiliations. We seek to reduce the importance of tracking individuals and track institutions instead. Currently, many data providers refer to user support offices as generic contacts rather then individuals. However, many individuals (data providers and investigators) still desire to be represented in the GCMD so that proper citation and credit for the data are given. We continue to track individuals as requested, but will refer to institution names whenever appropriate and applicable.

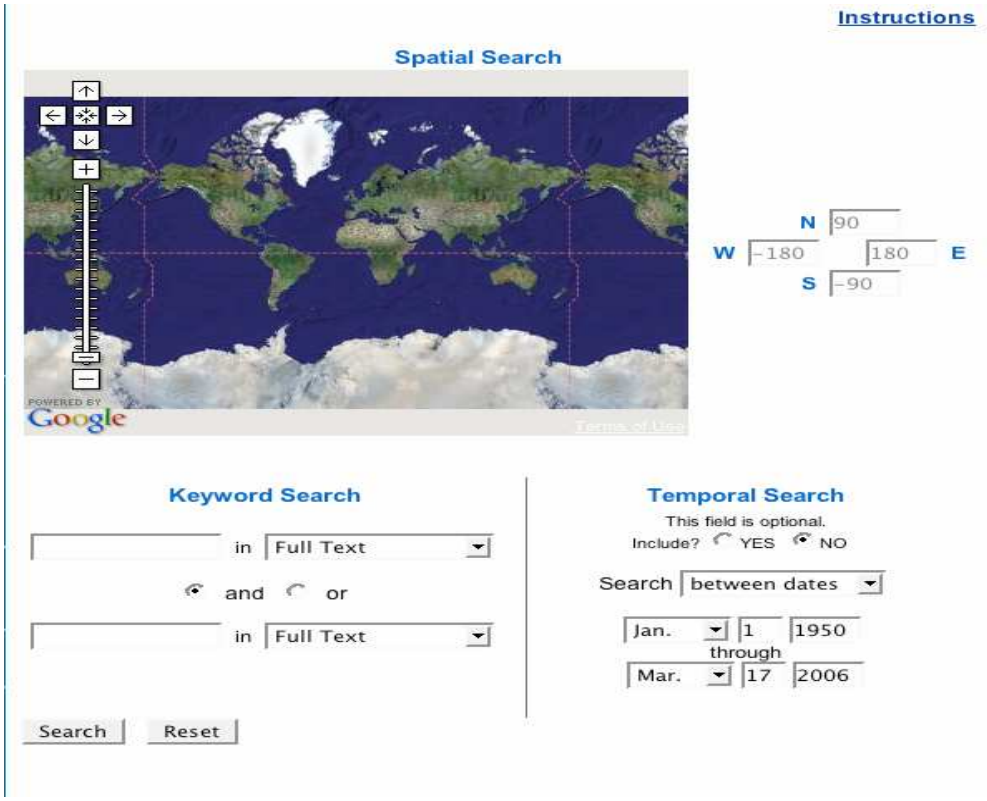


Data provider prefers personal contact

Many data providers refer to user support offices, rather than individuals

3. Searching by lat/lon and polygons.

The UWG recommended that support for polygon searches should have low priority. Therefore, we have instead focused on the geospatial search by incorporating the Google(r) map to improve the spatial search results. Users may combine the Google(r) map with free-text or fielded searching and, optionally, temporal search.



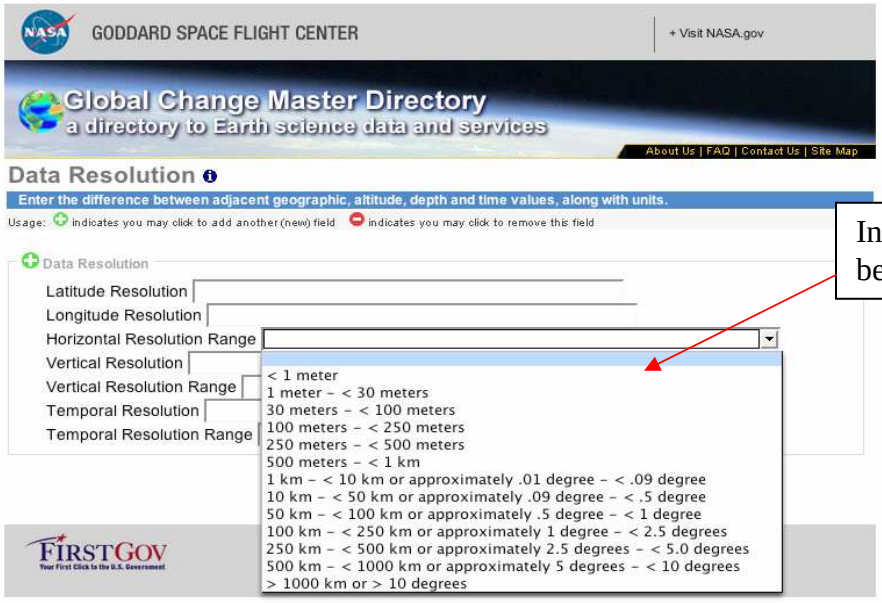
4. Refining spatial and temporal search capabilities.

A proposal to offer a spatial and temporal resolution search was submitted to the CEOS IDN interop and approved. Population of the resolution information in the NASA DIFs is proceeding with over 600 data sets populated with resolution information. Refinement of queries by resolution is now available for the GSFC, NSIDC, and LaRC DAAC portals. The key to usage is the appropriate set of controlled resolution keywords. The docBUILDER tools allow authors to populate the spatial and temporal resolution.

Users may specify the exact resolution and then provide the proper "bin" for that resolution for the search.

Horizontal Resolution	**Vertical Resolution**	**Temporal Resolution**
< 1 meter	< 1 meter	< 1 second
1 meter - < 30 meters	1 meter - < 10 meters	1 second -< 1 minute
30 meters -<100 meters	10 meters -< 30 meters	1 minute -< 1 hour
100 meters - < 1 km	30 meters -<100 meters	Hourly
1 km - <100 km or approx 0.1 degree -< 1 degree	100 meters -< 1km	Daily
100 km - <250 km or approx 1 degree -< 2.5 degrees	> 1 km	Weekly
Horizontal Resolution		**Temporal Resolution**
250 km - <500 km or approx 2.5 degree -< 5 degrees		Monthly
500 km - <1000 km or approx 5 degree -< 10 degrees		Hourly Climatology
1000 km or > 10 degrees		Daily Climatology
		Pentad Climatology
		Weekly Climatology
		Monthly Climatology
		Annual
		Annual Climatology
		Decadal
		Climate Normal (30-year climatology)

Sample screen shots show how to refine a query by temporal resolution buckets.

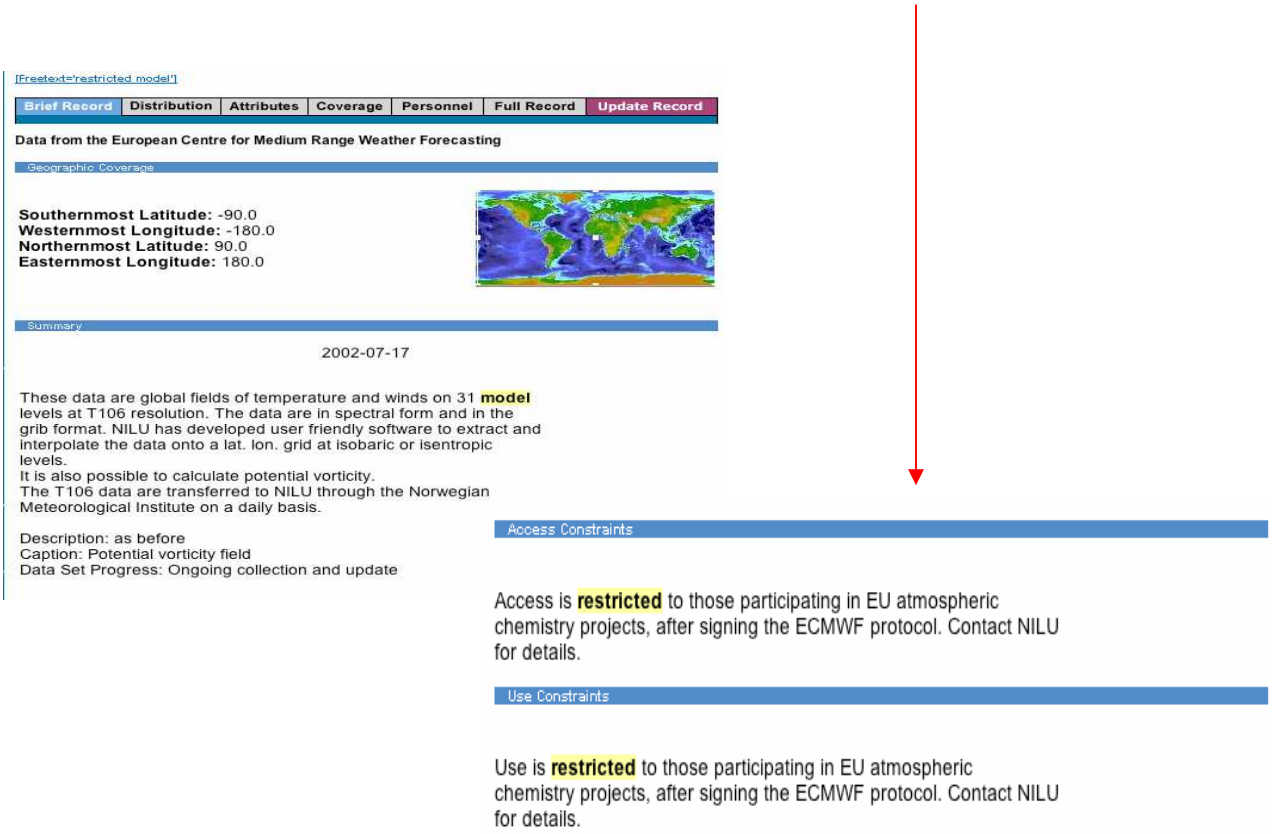


In docBUILDER, resolution buckets can be selected to include in metadata.

5. Private and restricted data access

a) Users need to be aware of model data that is available but restricted.

The DIF and SERF have fields for quantifying the access and use constraints of the model and/or model output. If model output is restricted, the information will be specified in these fields. Note highlights indicating the query terms, “model”, and “restricted” that were requested in the search.


[freetext='restricted model']

Brief Record | Distribution | Attributes | Coverage | Personnel | Full Record | Update Record

Data from the European Centre for Medium Range Weather Forecasting

Geographic Coverage

Southernmost Latitude: -90.0
Westernmost Longitude: -180.0
Northernmost Latitude: 90.0
Easternmost Longitude: 180.0

Summary

2002-07-17

These data are global fields of temperature and winds on 31 **model** levels at T106 resolution. The data are in spectral form and in the grib format. NILU has developed user friendly software to extract and interpolate the data onto a lat. lon. grid at isobaric or isentropic levels.
It is also possible to calculate potential vorticity.
The T106 data are transferred to NILU through the Norwegian Meteorological Institute on a daily basis.

Description: as before
Caption: Potential vorticity field
Data Set Progress: Ongoing collection and update

Access Constraints

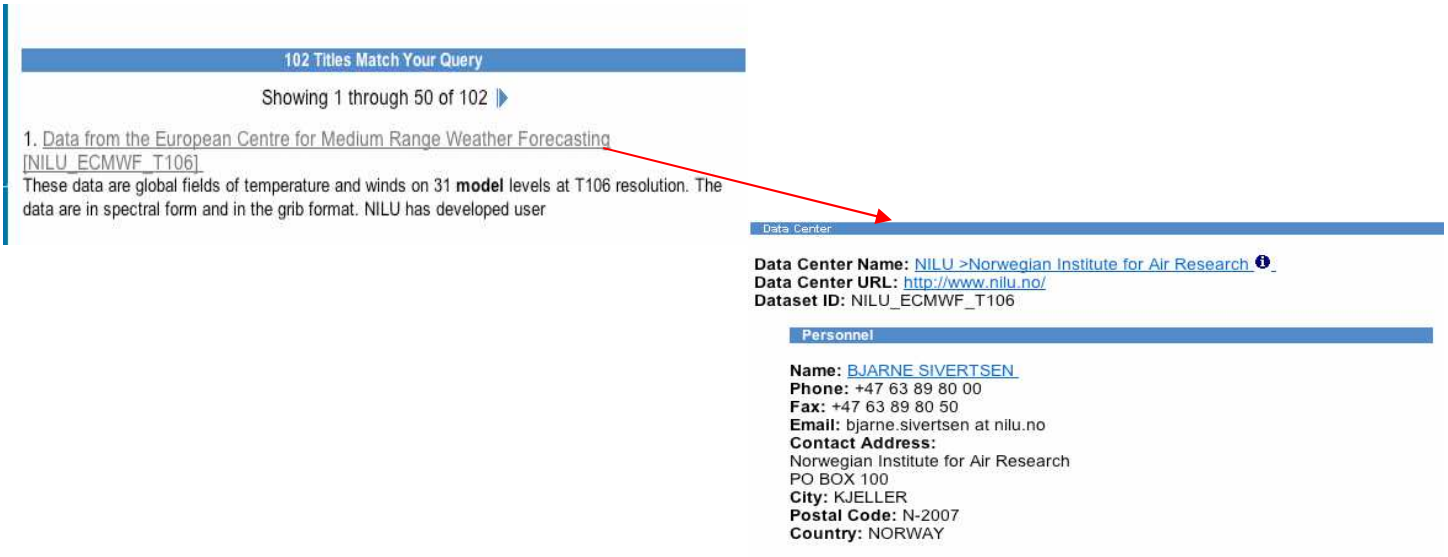
Access is **restricted** to those participating in EU atmospheric chemistry projects, after signing the ECMWF protocol. Contact NILU for details.

Use Constraints

Use is **restricted** to those participating in EU atmospheric chemistry projects, after signing the ECMWF protocol. Contact NILU for details.

b) Explore cataloging and facilitating the use of restricted data

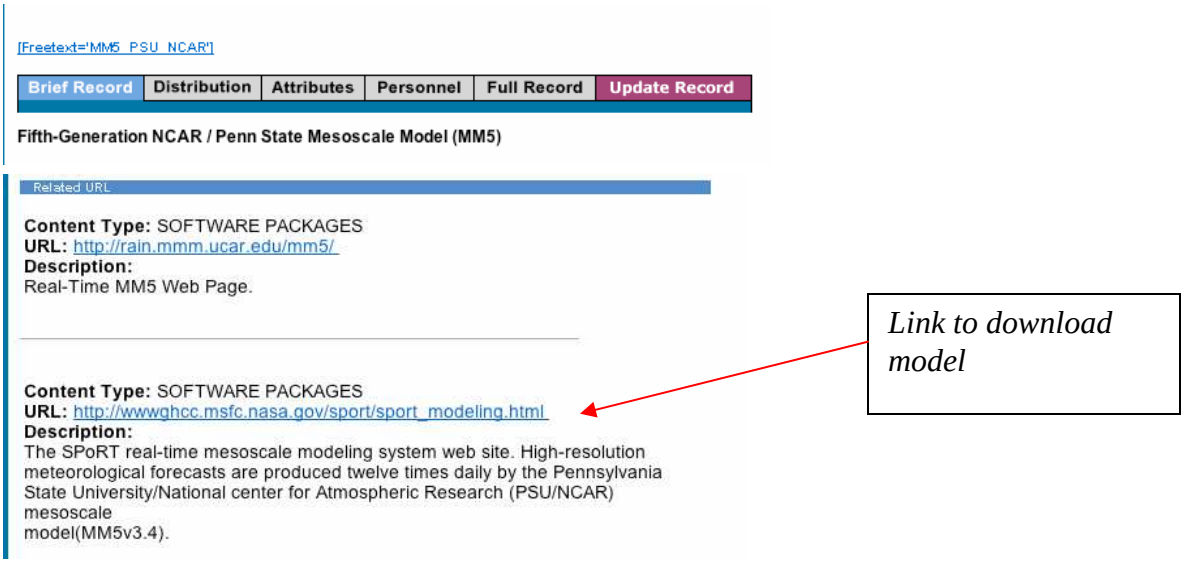
For restricted access data, the contact is provided:

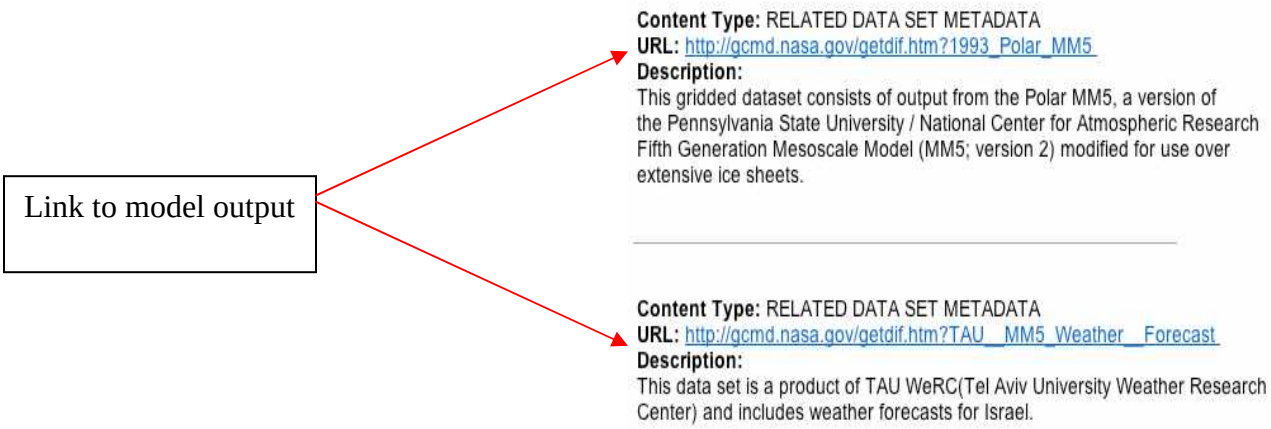


6. Provide better access to descriptions of scientific models

a) Include links and overviews for SERFs pertaining to models.

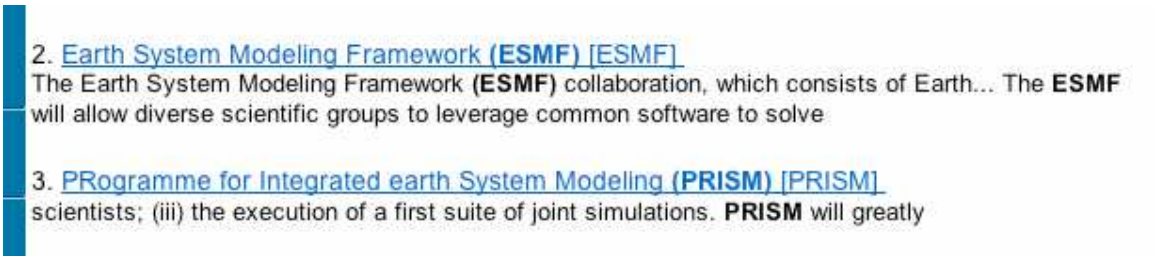
The SERF may provide online links to associated model output and offer links to directly access and download the model itself.





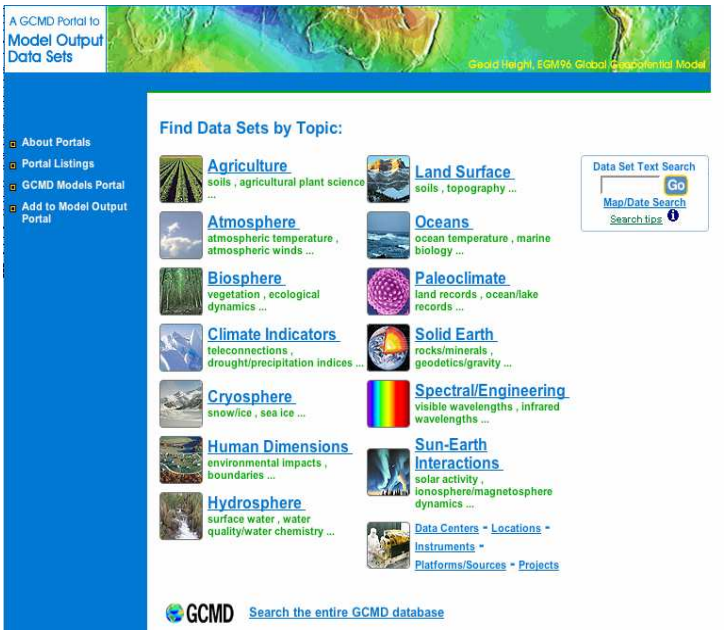
b) Include ESMF and PRISM model collections as SERFs.

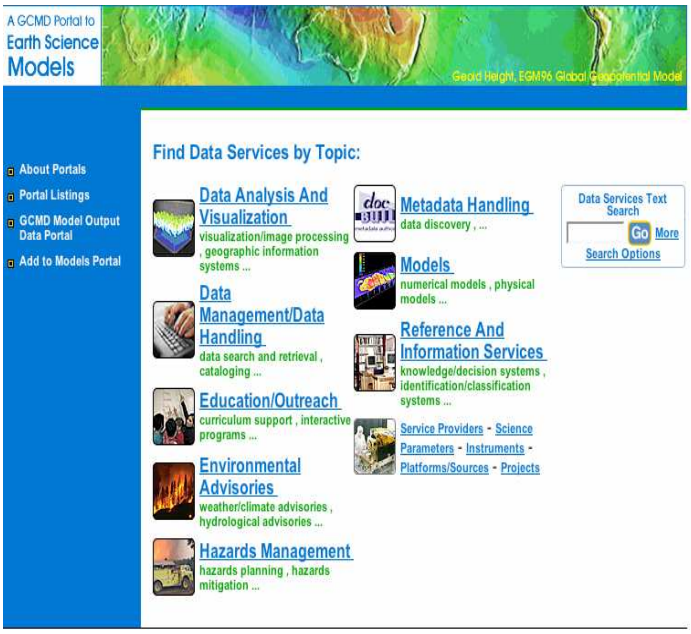
The Services Entry Resource Format (SERF) descriptions exist for the Programme for Integrated Earth System Modelling (PRISM) and the Earth System Modeling Framework (ESMF). If models or model output is submitted to the GCMD using these frameworks, we will add an appropriate keyword or Project identifier so that the descriptions can be located easily (e.g., Project: ESMF).



As of March 2006, 272 numerical models are described in the Services (SERF) database, and more than 650 model output data sets are described in the GCMD (DIF) database.

Portals have been implemented for models and model output:





c) Consider a workshop with modelers to better understand how models and results can be distributed.

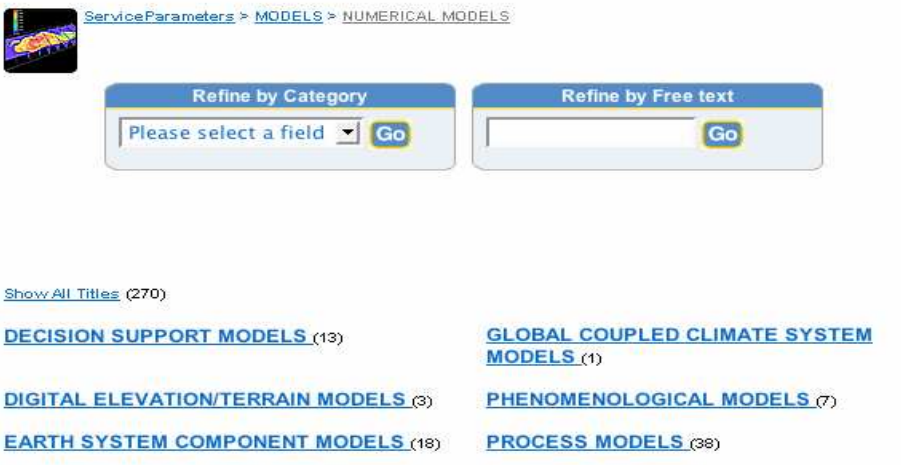
Modelers now participate in workshops under the auspices of the Global Organization for Earth System Science Portal (GO-ESSP) community. GCMD staff presented a paper at the 4th GO-ESSP workshop held June 6-8th, 2005 at the British Atmospheric Data Center (BADC), Rutherford Appleton Laboratory (CCLRC), Chilton, Didcot, England. A staff member was involved in discussions during break-out sessions addressing issues such as:
(1) CF standard names/standard vocabularies
(2) Proposed numerical DIF fields (i.e extensions to DIF)
(3) Data and metadata discovery (examples of discovery portals were presented - such as the NDG discovery repository, http://ndg.nerc.ac.uk/)
(4) OAI harvesting capabilities/protocol to follow among the GO-ESSP community (unique metadata identifiers, re-publishing of metadata.)
(5) Metadata profiles (DIF, WMO Core profile, ISO 19115)
(6) Better linkage to data (i.e., it was proposed that THREDDS capabilities should be imported in the GCMD DIF)
(7) Data and metadata standards (CF, NetCDF, HDF, THREDDS, MOLES, DIF).

A proposal by British Atmospheric Data Centre (BADC) is being drafted to extend the DIF to include information on numerical simulations (i.e. models and model output and referred to as "NumSim"). See: http://cgam.nerc.ac.uk/NMM/ipcc/index.php.

d) Describe models as "numeric" instead of "mathematical".

The TERM keyword has been changed to "Numeric Models" as suggested by the UWG. An experimental set of model keywords has been added at the VARIABLE level:

- Global Coupled Climate System Models *(Typically, atmosphere ocean, land, surface/hydrology/vegetation, snow/ice, atmospheric chemistry and radiation)*
- Global General Circulation Models *(Atmospheric GCMs, Ocean GCMs)*
- Earth System Component Models *(atmosphere, ocean, land, surface processes, hydrology, snow/ice, biospheric (vegetation, ocean biology), chemistry (atmosphere, ocean), radiation (chemistry and aerosols))*
- Process Models *(clouds and precipitation dynamics, cloud-radiation feedback, cloud-aerosol interaction, atmosphere-land surface coupling, atmosphere-ocean surface coupling, ocean-ice coupling/interaction, land hydrology (run-off models=P.E. + percolation), ecosystem models, transport models (pollution, trace constituents))*
- Phenomenological Models *(hurricane track/intensity prediction, El-Nino (ENSO) prediction, tsunami models, air quality index predictions, UV index predictions, heat wave predictions, drought index and predictions, flood warnings/predictions, disease outbreak warning systems/predictions (Bird flu, SARS, others)*
- Digital Elevation/Terrain Models
- Solar-Atmosphere/Space Weather Models (*typically, Solar-Earth atmosphere interaction warnings/prediction systems. Important for telecommunications, satellites, space operations, solar particle flux).*
- Decision Support Models



7. GCMD Calendar Service

a) Check and compare calendar service with AGU. Possibly share calendar entries.

Links are provided to 17 other calendar services including the American Geophysical Union: http://gcmd.nasa.gov/Resources/calendar/other_calendars.html. There is some overlap with events listed in the AGU calendar service; however, for the period March 20-April 2006, the AGU listed 7 events, while the GCMD/IDN calendar listed 17 events. We encourage those who request to advertise an event to use our calendar entry tool.

b) Track calendar hits.

At the end of 2005, there were 1,476 conference listings. Calendar hits are tracked on a monthly basis, as well as stats related to GCMD staff conference submissions and non-GCMD staff submissions. Last year (2005) the GCMD staff added 224 new conferences, updated 15 conferences and deleted 4 entries. These numbers are higher than for the previous year (2004) when 192 new conferences were added, 15 were updated, and 1 entry was deleted. GCMD partners added 10 unsolicited calendar entries in 2005. A response is always sent back to users to acknowledge their contribution. We also encourage use of the conference calendar as an easy, quick way to advertise a conference.

8. Copyright of metadata

The UWG suggested that the GCMD might wish to establish copyright of its metadata. We think that in lieu of copyrighting metadata, we should explore the standardization of uniqueness (beyond the unique "Entry_ID" field in the DIF/SERF). We will explore the implications of integrating an emerging standard for the publication and citation of scientific data being developed by the German Science Foundation, which is called the Scientific and Technical Data – Digital Object Identifier (STD-DOI). The system, which is operational for several World Data Center's in Germany, makes use of the DOI and URN infrastructure to uniquely represent and cite data sets.

9. Open Source software and migrating to Postsql

a) Continue using open source.

GCMD has incorporated open source software in operational and development environments. An internal product tracker database is kept on the Zope/Plone database for tracking open source and COTS software used on the GCMD project.

Examples of open source software packages used include:

Xplanner	project and task planning
TRAC	tracking bugs and enhancements
Lucene	free text search engine
Ant	Java-based build tool
Apache	web hosting software
Tomcat	web server used with Apache
Linux	a free Unix-based operating system
Subversion	source code control similar to CVS
Jython	Python for Java for scripting web-baeed servlets
Zope/Plone	web application server (used for internal documentation and IDN web site)
Squid	web component for proxying and caching web page
Struts	Java-based framework under Apache
Rsync	Unix file transfer program
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting

b) Explore process of migrating away from Oracle to postgreSQL.

NASA currently has a Goddard-wide agreement with Oracle, making the migration to another RDBMS less critical at this time. Testing with PostgreSQL has indicated questionable performance results; however, “hooks” to this database have been added to the latest MD9.4 package.

10. Upgrade web pages every 3 years.

In Fall 2004, NASA presented a new "One NASA" plan, which promoted creating consistent, identifiable NASA websites. The NASA GSFC webmaster site provided tools (e.g. graphics generator), information, and templates for converting websites to the "One NASA" design. Other NASA standards and guidelines were addresses by the GSFC Webmaster group. As of October 2005, a new *Section 508 of the Rehabilitation Act of 1973*, as amended in 1998, accessibility plan for reviewing NASA sites was adopted by GSFC. The Section 508 web accessibility compliance standard (*GPR 2800.1, GSFC Section 508 Web Compliance*) addresses the importance of providing both employees and members of the public with disabilities access to websites. Additional information about 508 compliance guidelines is available online: http://web508.gsfc.nasa.gov/.

As new NASA guidelines and standards are presented to the GSFC webmasters group, GCMD will adhere to NASA policies and best practices and makes changes to the site as required. We incorporated the "One NASA" design in March 2005 (release of MD9.4 software) and are continually updating content on the site to meet 508 compliance guidelines.

11. Fix dead links.

Identifying and fixing broken links is an on-going process and a maintenance burden. We currently use the *W3C Link Checker Version, 4.0* software. Approximately 150 links/month are identified and fixed. According to the "NASA First" policy, an intensive effort has been deployed to ensure links to NASA data and resources are operational. Our strategy is to complete one full scan of the metadata collection to identify broken links every 12 months.
Total: 16,480 DIFs
Total: 1,379 SERFs

"Link Checking Activity (2005-01-05 - 2006-02-28)"

CATEGORY	# OF DIFS	DATE SCANNED FOR BROKEN LINKS	DATE & CONTINUED STATUS OF LINKS
ACADEMIC	2,219	12/2/2005	1/12/2006 A. Ajayi Completed Fixing Links
COMMERCIAL	265	5/24/2005	5/27/2005 A. Ajayi Completed Fixing Links
CONSORTIA/INSTITUTIONS	442	2/9/2005	2/28/2006 Links Not Checked Yet

DOC	1,891	2/8/2005	3/16/2005 S. Ritz Completed Fixing Links
DOD	101	2/8/2005	2/28/2006 Links Not Checked Yet
DOE	244	2/25/2005	2/28/2006 Links Not Checked Yet
DOI	1,198	2/8/2005	2/28/2006 Links Not Checked Yet
EPA	107	2/13/2006	2/13/2006 A. Ajayi Fixing Broken Links
GOV. AGENCIES, NON-US	3,883	2/25/2005	2/28/2006 Links Not Checked Yet
MULTINATIONAL ORGS.	1,910	6/6/2005	2/28/2006 Links Not Checked Yet
NASA-EOSDIS	2,623	12/14/2005	12/14/05-2/27/06 S. Ritz fixed links in 70 DIFs; added GET DATA link
GES-DAAC	456	4/20/2005	4/1/2005 Majoe reconstructuring of GES DAAC web tree
			4/20-08/31/2005 S. Ritz revised links in all DIFs to link to web tree (400 difs)
LARC DAAC	387	7/26/2005	07/27/2005 Sent broken link report to K. Morris
ORNL DAAC	767	10/26/2005	10/30/2005 S. Ritz Completed Fixing Links
ASF DAAC	33	10/26/2005	10/30/2005 S.Ritz Completed Fixing Links
MSFC/GHRC	127	10/26/2005	10/30/2005 S. Ritz Completed Fixing Links
NON-PROFIT ORGS.	161	5/27/2005	05/31/2005 A. Ajayi Completed Fixing Links
NSF	1,032	3/3/2005	02/28/2006 Links Not Checked Yet
US STATE/REGIONAL/LOCAL AGENCIES	335	NOT SCANNED YET 2/28/2006	
USDA	652	2/13/2006	02/14/2006 A. Ajayi Completed Fixing Links
SERFS		7/19/2005	7/19/2005 Report to H. Weir

12. Adopt a policy for retiring old DIFs.

If a dataset is no longer available or the author cannot be contacted, the data set record will be deleted. An automated reminder will be available for DIF authors whose records are more than 2 years old (see Section 8c).

Metadata records can be archived under the same versioning software that is used for source code control (Subversion). Steps are being taken to:

1. Create initial db of all records.
2. Modify loader to perform a “commit” on a changed record.
3. Set up web interface to show current record and revisions

13. Organization of technical detail of the UWG meetings

The UWG has recommended that future meetings be divided into 2 parts. Day 1 should be system configuration and technical issues for invited and interested participants and Day 2 would be higher-level programmatic issues. We agree with the UWG recommendations and will organize the presentation of information as suggested for future UWG meetings.