

The State of the GCMD - CY 2003

I. Mission

The mission of the Global Change Master Directory (GCMD) is to assist the scientific community in the discovery of and linkage to Earth science data and related services, as well as to provide data holders a means of advertising their data to the Earth Science Community. The GCMD leads the search to data and offers links to data providers and often directly to data.

II. Organization

The directory is staffed by three software developers, a systems administrator, a database administrator, and four Earth science coordinators (one coordinator is supported by the USGS Biological Resources Division), the contract task lead, and the project manager, Lola Olsen. The science coordinators are responsible for the data sets related to the Biosphere and Land Surface; Oceans and the Hydrosphere; Solid Earth and the Cryosphere; Solar-Terrestrial Interactions; Spectral/Engineering data; Agriculture and Human Dimensions; the Atmosphere and Climate Indicators; and Paleoclimate. The software developers collectively hold skills in database management, programming languages such as Python, Java, Perl, and C++, user interface design, and configuration management.

III. Science User Working Group

Representing the broad range of Earth science disciplines including life sciences, oceanography, geophysics, and atmospheric science is an active Science User Working Group. Members are chosen for their interest in NASA's directory effort, and for their understanding of the importance of high quality data management. The composition of the group for the upcoming UWG meeting in May 2004 includes Acting Chairperson, Mr. Martin Ruzek (USRA); Dr. Walter R. Hoegy, NASA/Goddard representative from Code 910; Dr. Erick Chiang, National Science Foundation/Office of Polar Programs; Dr. Glenn Rutledge, NOAA/National Climatic Data Center (NCDC); Dr. Benno Blumenthal, Lamont-Doherty Earth Observatory at Columbia University; Dr. Wendell Brown, University of Massachusetts-Dartmouth; Dr. Hubert Staudigel, Scripps Institution of Oceanography; Ms. Andrea Buffam from the Canadian Centre for Remote Sensing (CCRS), representing the CEOS IDN; and Dr. Doug Beard, USGS/Biological Resources Division (BRD). The GCMD Science User Working Group works under the UWG Terms of Reference. Recommendations from previous UWG meetings have helped to guide the progress of the directory.

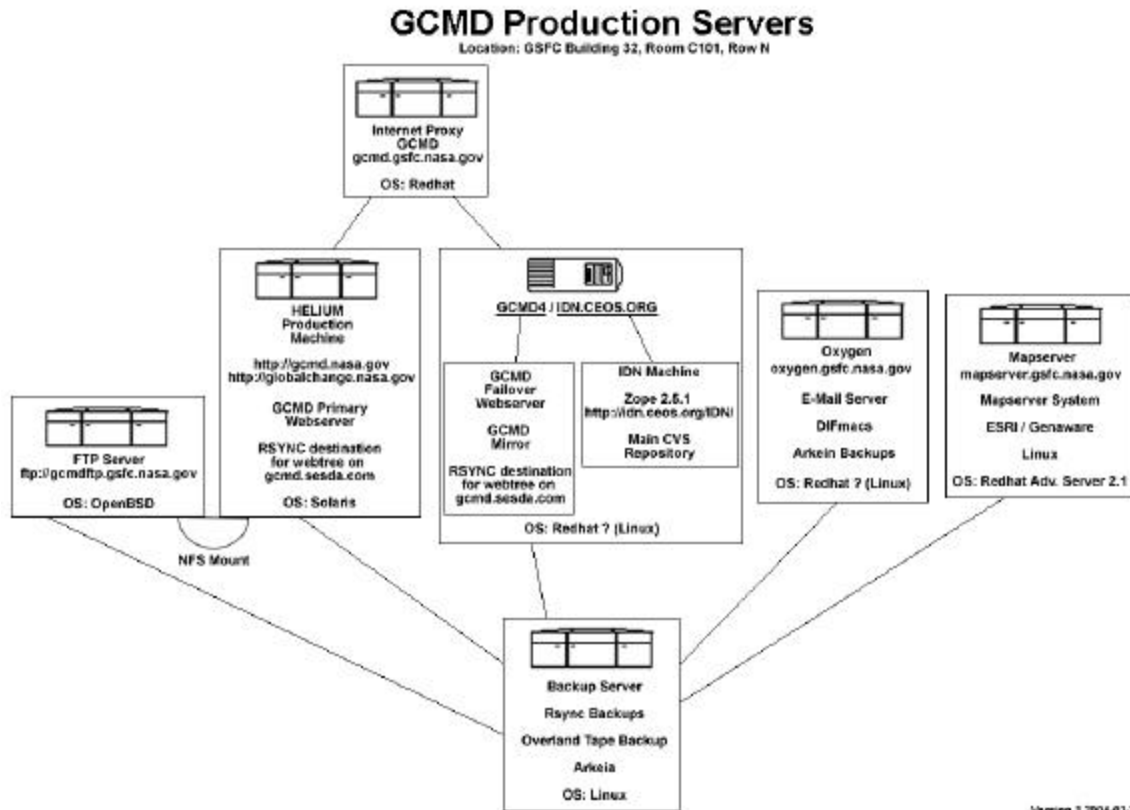
IV. Highlights for 2003

The top five highlights for 2003 that have had the greatest positive impact on the GCMD work include:

1. *Improved configuration management, with expanded testing through a set-up of “sand-box” machines.* Extensive test plans to complete functional end-to-end testing of all major system components on a variety of platforms were designed for quality assurance.
2. *The recognition of the value of the “portals” by an increasing number of interested groups - virtual subsets of the directory that can be created for special events and/or for groups wishing to host a directory of only their contributions.* During 2003, eleven new portals were created.
3. *The addition of an “Update” option within the authoring tools, permitting updates to be made quickly and easily by external participants.* This important advancement provides an important key to maintaining current information in the directory.
4. *A 47.7% rise in content for Earth science tools and services - creating increased interest in fruitful searches and the impetus to add new content.*
5. *The addition of a new topic keyword category, “Climate Indicators”, which is now the most frequently searched of the 13 topics.*

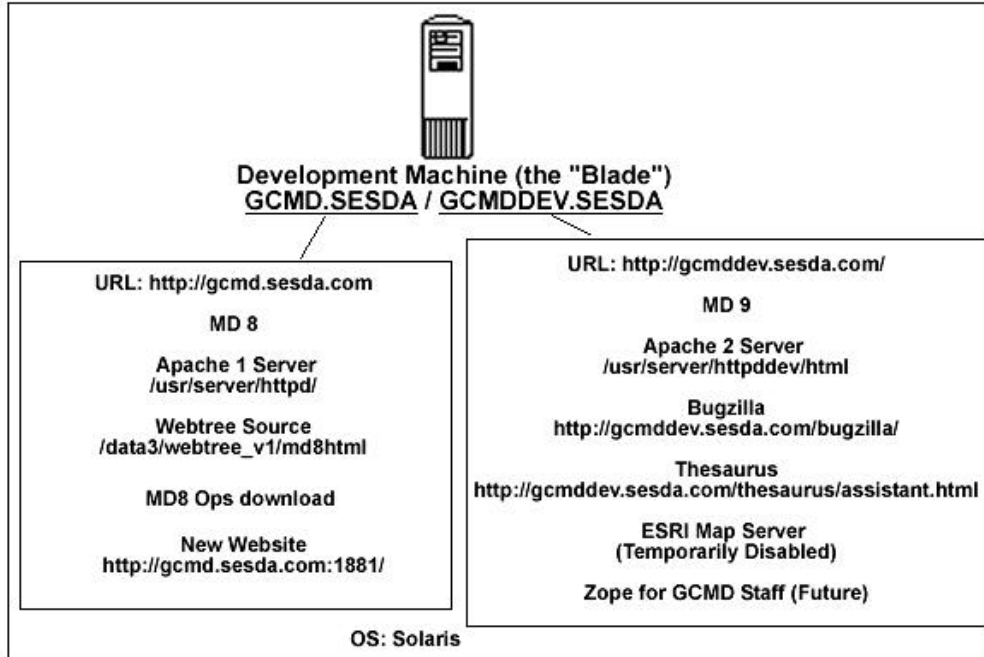
V. Facility for Operations and Maintenance

The GCMD facility is located at the Goddard Space Flight Center. The primary operational computing resources consist of a “proxy” computer and a backup for the proxy - accessible as <http://globalchange.nasa.gov>. These computers route users to one of two operational computers, so that one is always available for user access when the other needs maintenance or requires the installation of upgrades. To ease the network burden for off-site contractors, a “development platform” is available at the contractor’s facility. New this year are a series of testing platforms (several acquired through Goddard’s equipment “excess”) to perform cross-platform testing and quality assurance. See diagrams.



GCMD Development Server

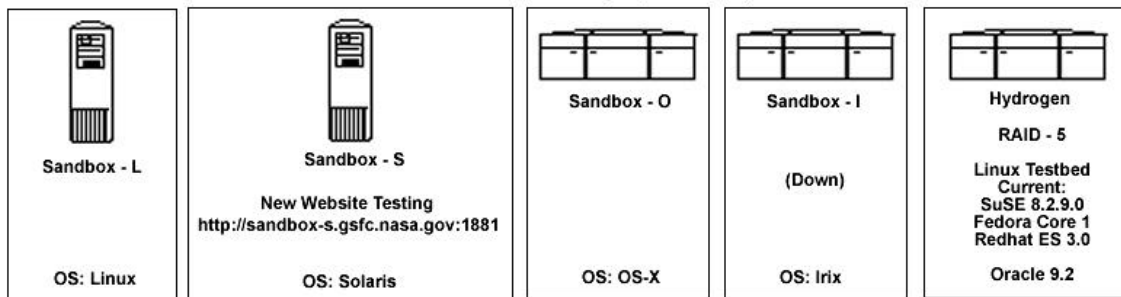
Location: Aerospace Building, 5th floor, room 564



Version 1 2004-01-30

Development/Testing Sandboxes

Location: GSFC Building 32, Room C101, Row J



Version 1 2004-01-30

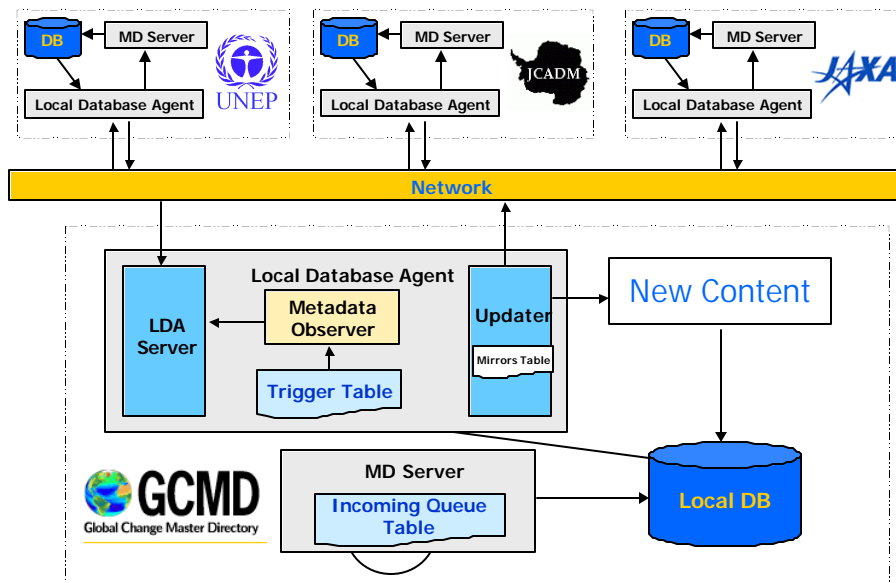
VI. System Development

To make electronic and information technology accessible to people with disabilities, Congress amended the Rehabilitation Act. The law is known as “508”, and the GCMD’s MD8 software is fully “508” compliant.

Software upgrades were integrated (transparently to the users) during the year – as the code was re-factored “behind the scenes” into a better-organized, more reusable package in preparation for the ISO-19115 implementation. Users may note visible upgrades such as (1) a new capability to refine the current query to include global, global land, and global ocean data sets; (2) an upgraded parent/child functionality that allows users to refine by children only, parents only, and navigate more easily between the two; and (3) more direct access to data such as MODIS through links located next to the titles’ display.

For the first time last year, The Operations Facility (OPS) was made available to all the science coordinators to interact directly with the database. The facility performed well in 2003. Changes were also made within the Local Database Agent (LDA) software for improved functionality of the distributed aspect of the system. See revised diagram.

Local Database Agent



In addition, the Geospatial One-Stop was completed, permitting users to access information easily by project. This search for NASA projects (missions) was based on the development of a free-text search through the open-source search engine, Lucene.

See: <http://gcmd.gsfc.nasa.gov/md/lucene/luceneSearch.html>

The work was done in response to OMB A-11 Guidance regarding posting information on geospatial data acquisitions and on compliance with FGDC standards. A-11, Section 51 Basic Justification Materials: In February 2003, agencies must post information on all geospatial data acquisitions in excess of \$1M planned for FY 2004 on the FGDC Clearinghouse, characterized using the FGDC metadata standard, taking care to specify the geographic area and scale to which the data layer is proposed to be collected. Exhibit 300 Section II.A.2 Data C. If this initiative processes spatial data, identify planned investments for spatial data and demonstrate how the agency ensures compliance with FGDC standards required by OMB Circular A-16. See the web site, <http://www.geo-one-stop.gov/>

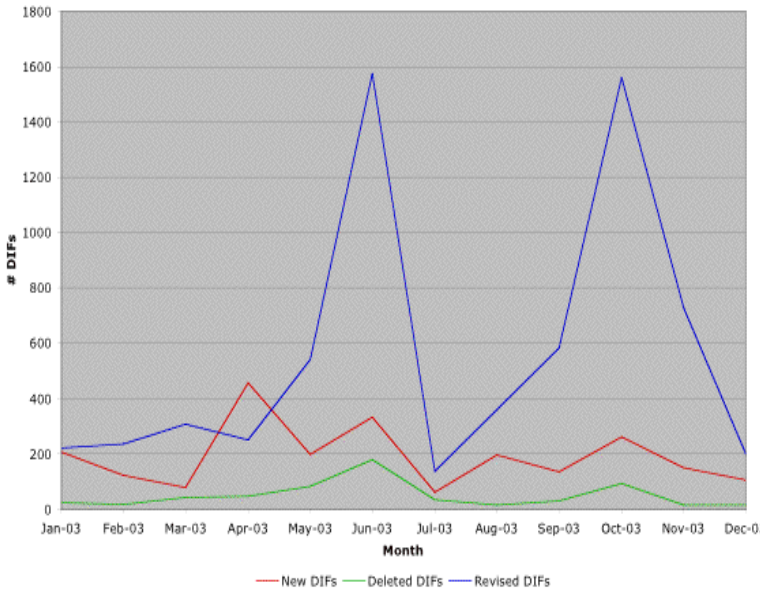
VII. Metrics for Content Usage and User Feedback/Support

A. Content

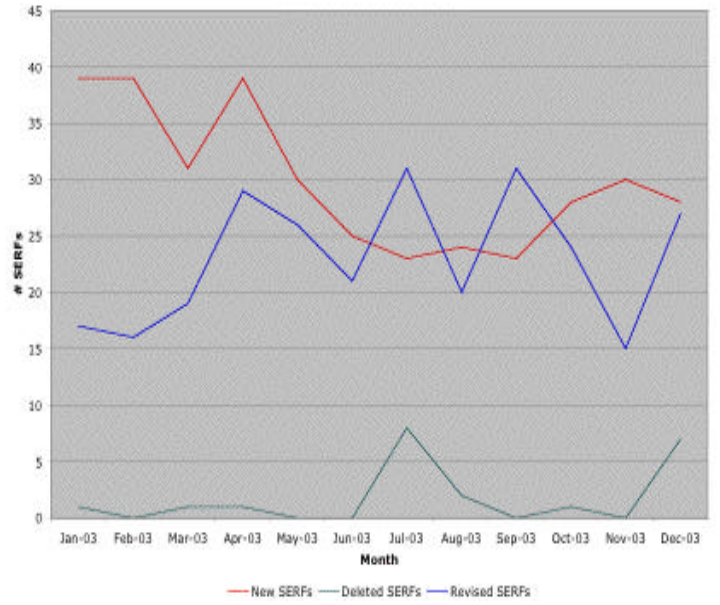
The population of the directory's data sets reached 14,049 at the end of December 2003, up from 12,343 the previous year. The number of new data set descriptions decreased from 2,503 in CY02 to 2293 in CY03; however, the number of revised data set descriptions increased 4.9%, from 6,376 in CY02 to 6,691 in CY03. The distribution of DIFs among Earth science topics can be viewed in the chart below:

GCMD POPULATION	2002	2003	% CHANGE in 2003 input.
New DIFs	2503	2293	-8.3%
Revised DIFs	6375	6691	+4.9%
New SERFs	243	359	+47.7%
Revised SERFs	213	276	+29.6%
New Data Centers	42	333	+692.8%
Revised Data Centers	27	21	-22.2%
New Sources	52	144	+176.9%
Revised Sources	12	69	+475.0%
New Sensors	5	519	+10,280%
Revised Sensors	23	56	+143.5%
New Projects	401	84	-79.0%
Revised Projects	37	32	-13.5%

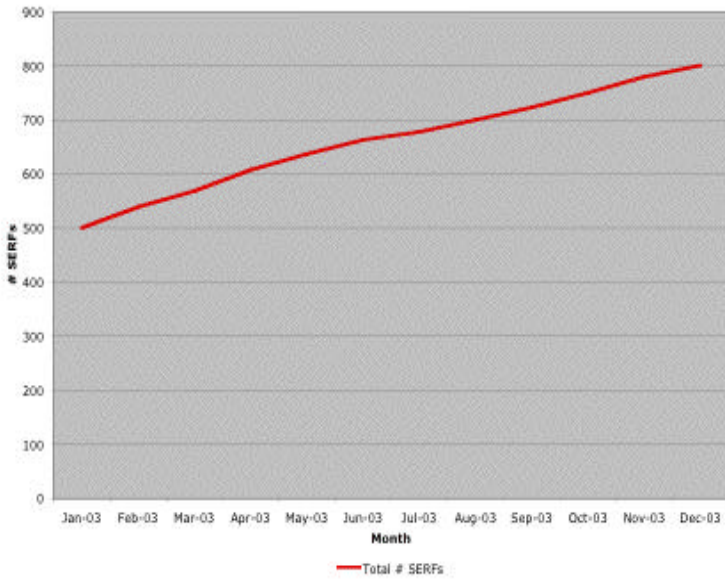
DIF Population 2003



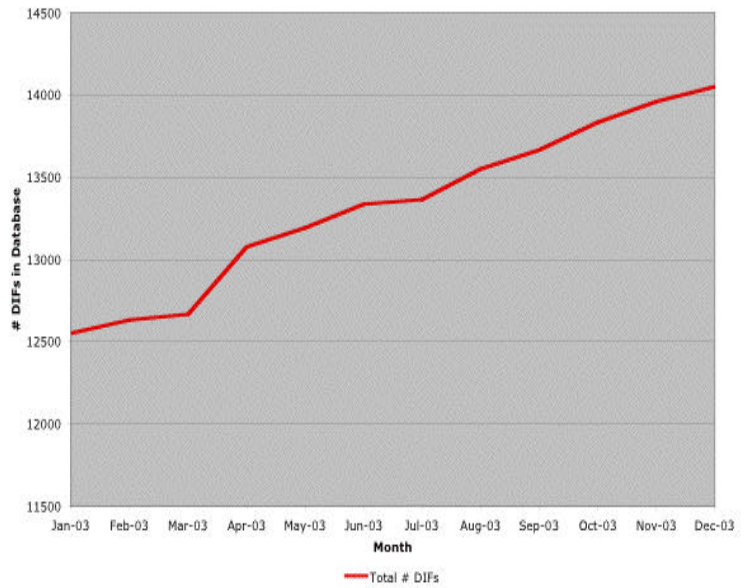
SERF Population 2003



Total # SERFs 2003



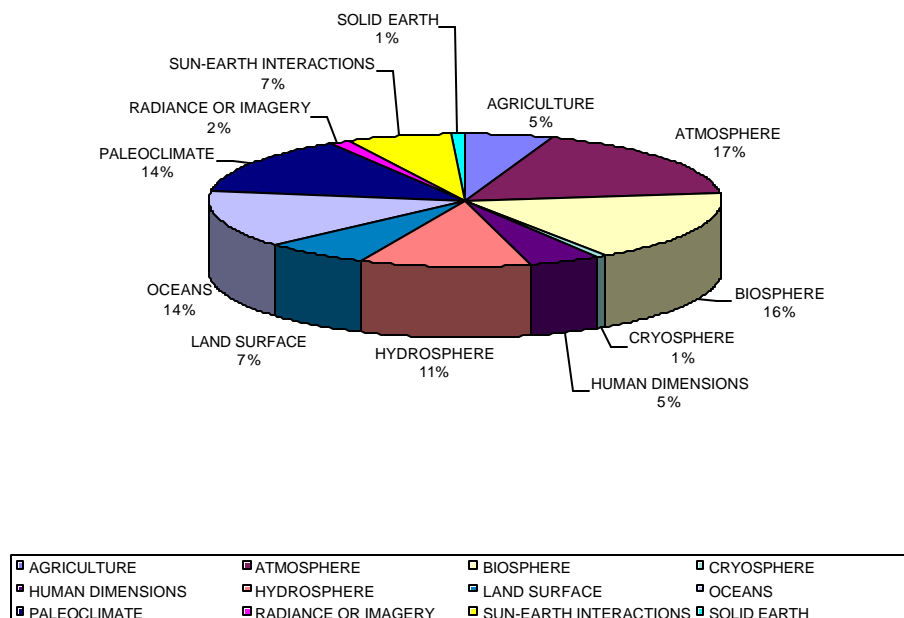
Total Number of DIFs 2003



Population of DIFs by Earth Science Topic

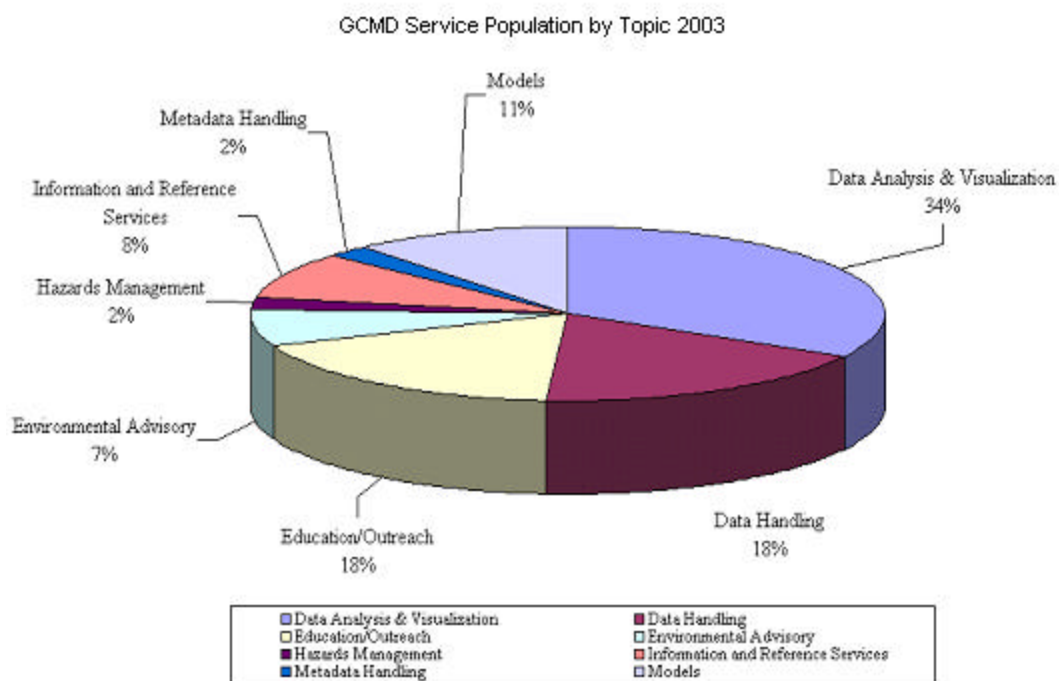
TOPIC	2002	2003	% CHANGE
AGRICULTURE	168	139	-17.3%
ATMOSPHERE	565	617	+9.2%
BIOSPHERE	696	798	+14.6
CLIMATE INDICATORS	N/A	21	N/A
CRYOSPHERE	336	218	-35.1%
HUMAN DIMENSIONS	629	398	-36.7%
HYDROSPHERE	245	332	+35.5%
LAND SURFACE	722	553	-23.45
OCEANS	415	702	+69.1%
PALEOCLIMATE	48	93	+93.7%
RADIANCE/IMAGERY	255	160	-37.2%
SUN-EARTH INTERACTIONS	6	13	+116.7%
SOLID EARTH	340	322	-5.3%

GCMD Population by Topic 2003



Population of SERFs by Earth Science Service Topic

Topic	2002	2003	% Change
Data Analysis & Visualization	199	374	87.94%
Data Handling	114	206	80.70%
Education/Outreach	139	199	43.17%
Environmental Advisory	46	78	69.57%
Hazards Management	11	28	154.55%
Information and Reference Services	51	96	88.24%
Metadata Handling	7	25	257.14%
Models	51	126	147.06%



The GCMD continues to offer a variety of metadata authoring tools for the creation and modification of DIFs and SERFs. The DIF writing tools are widely used among the EOSDIS Distributed Active Archive Centers (DAACs). LaRC/ASDC, NSIDC, GES, MSFC/GHRC, SEDAC, and PODAAC regularly send all their new and modified DIFs using the DIF authoring/modification tools. The DAACs have also contributed new and revised SERFs using the online tools. In total, the DAACs contributed 431 DIFs and 17 SERFs to the GCMD in 2003. Development of new, XML-based and Java-based metadata authoring tools (docBUILDER) is well underway, with the release scheduled for early 2004.

Several of the Federation's Earth Science Information Partners (ESIPs) have provided DIFs and SERFs using the metadata authoring tools. The University of New Hampshire (EOSWEBSTER), DODS/OPeNDAP, and the Museums Teaching Planet Earth (MuTPE) were the most active contributors during the year. In total, the non-DAAC Type 1, Type 2, and 3 ESIPs contributed 254 new DIFs and 25 new SERFs.

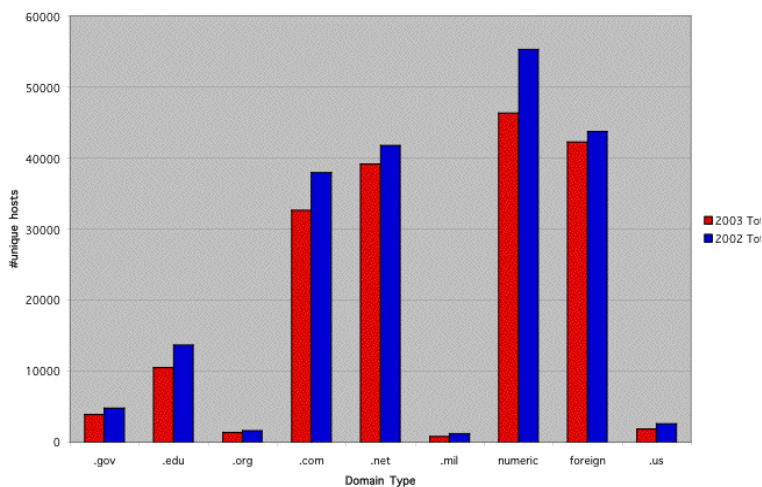
1. A critical aspect of the GCMD mission is that of providing links to data sets. At the end of CY03, the GCMD provided 70,535 links to data sets and information within the DIF and SERF metadata - up from 52,539 in CY02, and an additional 4,327 links from within supplemental descriptions (instruments, platforms, projects, data centers) and 1,430 links to web resources on the Earth Science pointers page.
2. The Global Change Conference Calendar now contains 1007 conference listings. During 2003, 202 new conferences were added and 17 were revised.

B. GCMD Web Usage

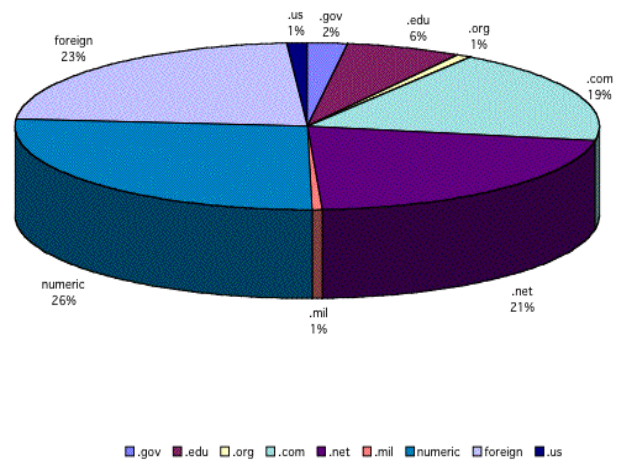
Usage, as measured by the number of unique hosts, for CY03 was down by 11.7% from 202,518 to 178,730 unique hosts – reflecting the inability of the metrics software to determine unique users among the collapsed “.com” addresses. The number of “hits” on the GCMD web site continued to climb from 6,463,527 in CY02 to 7,569,397 in CY03 (+17.1%). User web sessions show a remarkable “academic year”-like pattern of low activity during the summer and winter holidays.

U.S. UNIQUE HOSTS	2002	2003	%CHANGE
.gov	4,774	3,925	-17.8%
.edu	13,637	10,469	-23.2%
.org	1,625	1,332	-18.1%
.com	37,921	32,613	-14.0%
.net	41,793	39,135	-6.4%
.mil	1,207	831	-31.1%
.us	2,566	1,849	-27.9%
Numerical	55,263	46,351	-16.1%
HITS	6,463,527	7,569,397	+17.1%

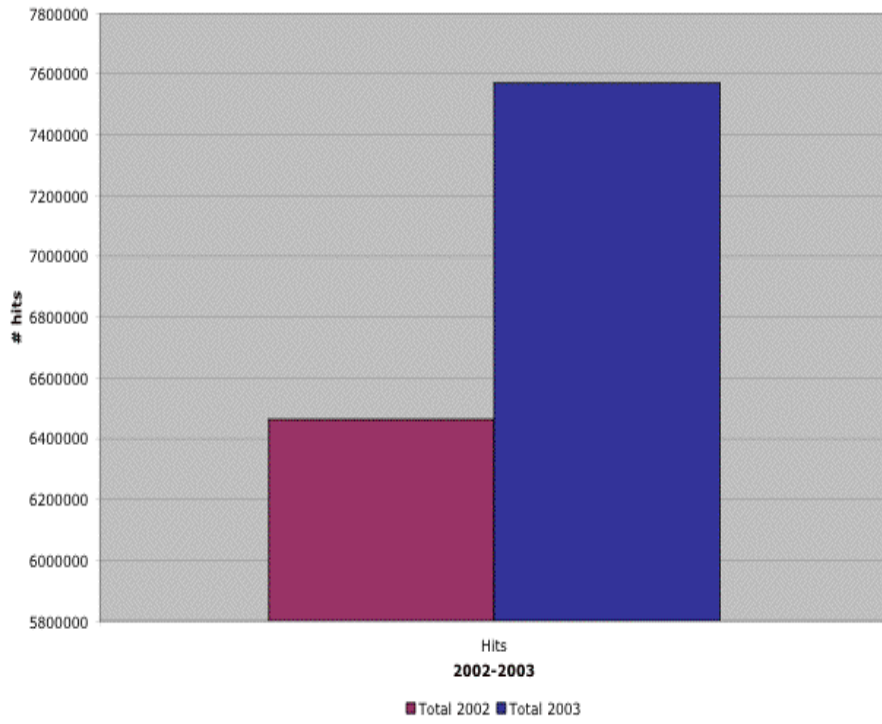
2002-2003 Web Usage by Domain



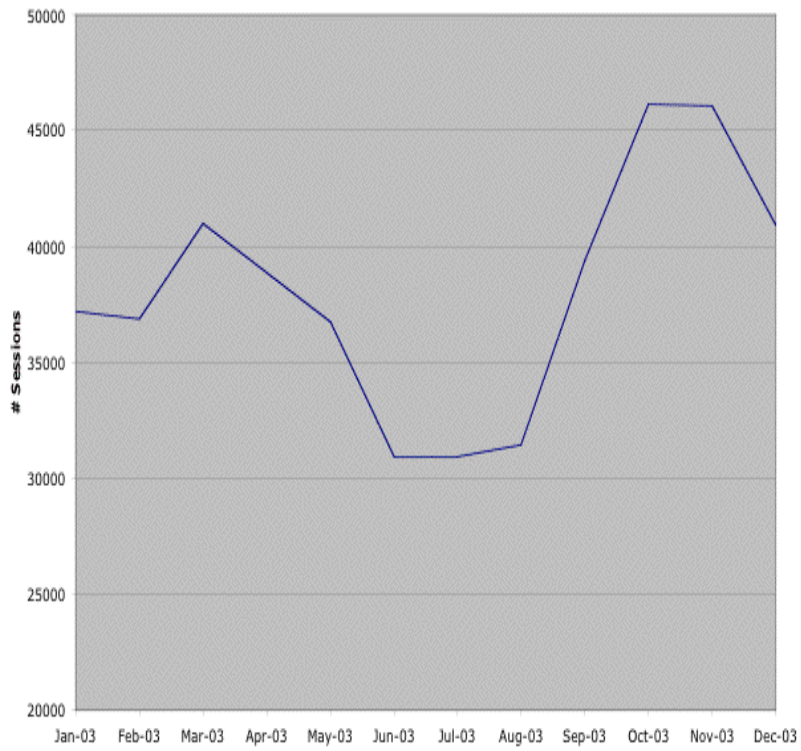
GCMD usage by domain type 2003



GCMD Web Hits 2002-2003

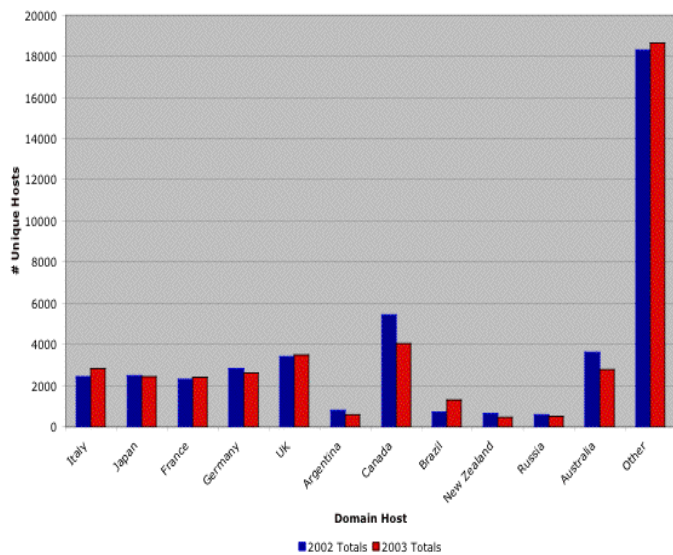


User Sessions 2003

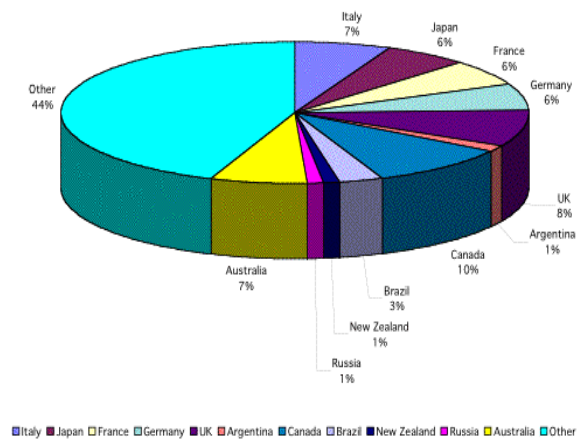


INTERNATIONAL HOSTS	2002	2003	%CHANGE
All International hosts	43,732	42,225	-3.4%
.it (Italy)	2,450	2,841	+15.9%
.jp (Japan)	2,508	2,451	-2.3%
.fr (France)	2,306	2,405	+4.3%
.de (Germany)	2,845	2,618	-8.0%
.uk (United Kingdom)	3,431	3,499	+2.0%
.ar (Argentina)	808	612	-24.0%
.ca (Canada)	5,446	4,063	-25.4%
.br (Brazil)	735	1,304	+77.4%
.nz (New Zealand)	653	469	-28.2%
.ru (Russia)	598	512	-14.4%
.au (Australia)	3,634	2,795	-23.1%
Other International	18,318	18,656	+1.8%
All Unique HOSTS	202,518	178,730	-11.7%

International Hosts 2002-2003 Difference



International User Domains 2003



Search and Retrieval from free-text and keyword interfaces

INTERFACE	2002	2003	%CHANGE
Search: Free- Text	51,559	57,730	+11.9%
Retrieval: Free-Text	26,138	25,912	-0.9%
Search: Zserver	50,261	204,216	+306.3%
Retrieval: Zserver	115,755	133,540	+15.3%
Retrieval: Supplemental	528,961	195,709	-63.0%

KEYWORD	2002	2003	%CHANGE
Search: Keyword	124,427	80,839	-35.0%
Retrieval: Keyword	66,644	257,211	+285.9%
Search: Portal Keywords	33,150	12,123	-63.4%
Retrieval: Portal Keywords	27,994	24,772	-11.5%

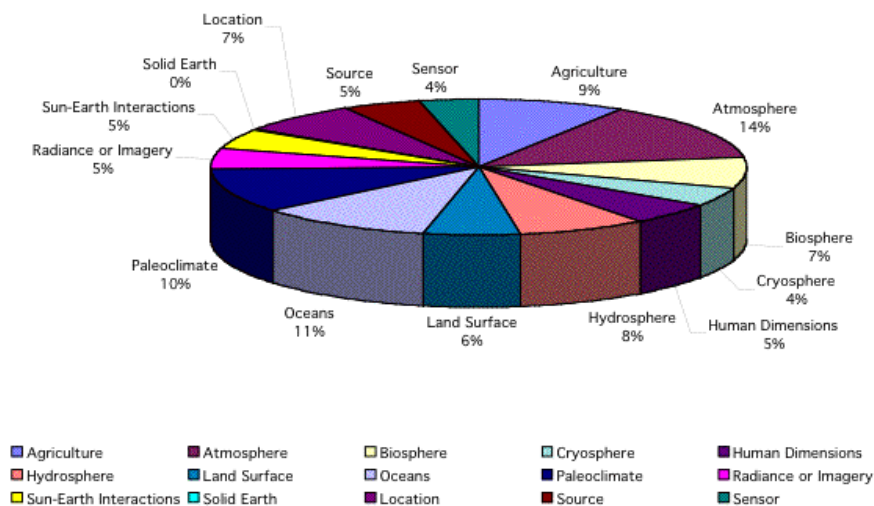
TOTAL	2002	2003	%CHANGE
Search: DIF	157,577	92,962	-41.0%
Retrieval: DIF	94,638	281,983	+197.9%
Search: SERF	8,578	10,684	+24.5%
Retrieval: SERF	42,432	63,011	+58.5%

NON-GCMD RETRIEVALS	2002	2003	%CHANGE
EDG	6,197	133	-97.8%
GLIS (USGS)	24,414	7,113	-70.9%
GSFC DAAC	3,004	31,399	+945.2%
OTHER	270,841	396,252	+46.3%

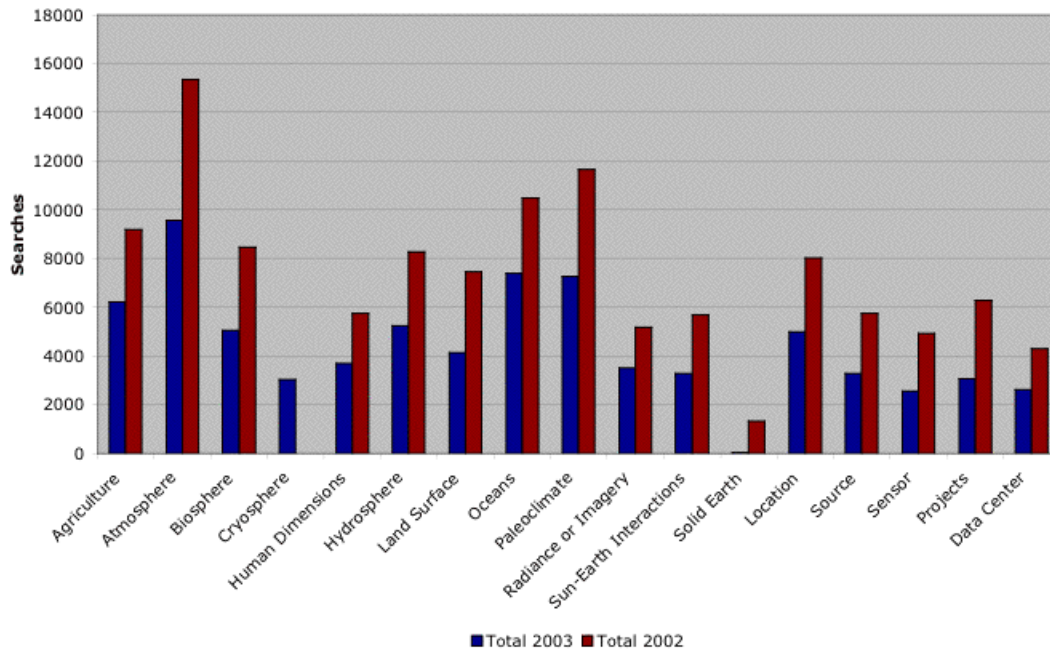
TOTAL RETRIEVALS	2002	2003	%CHANGE
Retrievals: ALL SOURCES	540,471	690,685	+27.8%

PARAMETER SEARCHES	SEARCHES CONDUCTED	%OF TOTAL SEARCHES
AGRICULTURE	6,203	7.7%
ATMOSPHERE	9,565	11.8%
BIOSPHERE	5,029	6.2%
CLIMATE INDICATORS	3,031	3.7%
CRYOSPHERE	3,695	4.6%
HUMAN DIMENSIONS	5,244	6.5%
HYDROSPHERE	4,118	5.1%
LAND SURFACE	7,731	9.1%
OCEANS	7,244	9.0%
PALEOCLIMATE	3,503	4.3%
RADIANCE/IMAGERY	3,281	4.1%
SUN-EARTH INTERACTIONS	26	0.03%
SOLID EARTH	4,987	6.2%
LOCATIONS	3,290	4.1%
PLATFORMS	2,558	3.2%
INSTRUMENTS	3,064	3.8%
PROJECTS	2,612	3.2%
DATA CENTERS	6,018	7.4%

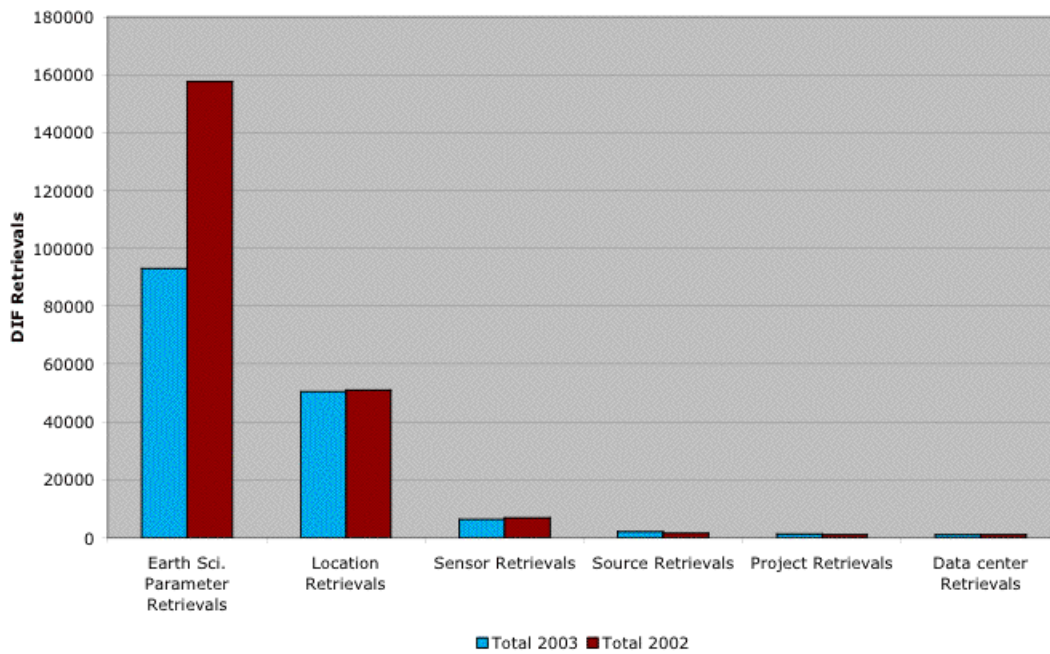
Searches by Earth Science Topic



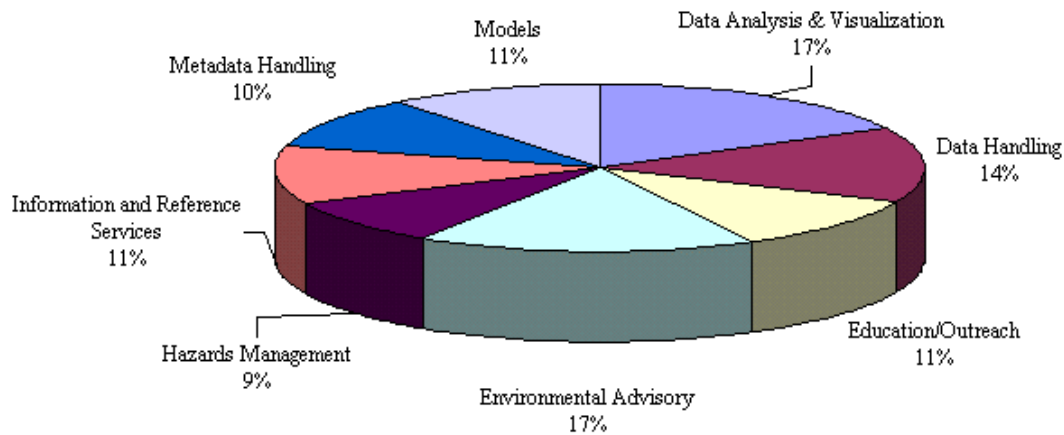
2002-2003 Keyword Searches



2002-2003 Retrievals by Keyword



**Searches by Service Parameter Topic
2003**



C. User Feedback and User Support

101 user support questions were received during 2003. Users typically ask questions about locating maps and images, climate data, and sources of information on various science topics. Below are some examples of user support questions received:

"I am writing a term paper about the Northwest Passage. I want to write about three major points; 1. The climate change, the melting of the ice, the lengthening of the melting periods 2. The legal status of the route, legal character 3. This route as a navigable route My question to you: can you help me finding information or could you provide me information?"

"I am looking for an aerial or satellite photograph of my father in-law's land in British Columbia, Canada."

"Could you please tell me the national map projection parameters for Bolivia?"

"I have EndNote6 (ISI) a software which allows downloading of records for reference purposes (scientific papers) if the service is Z39.50 compliant. I imagine that your server is. I do need information to set up a connection file for downloading. Can you help me?"

"I'm looking for raw data on global temperatures over the last say 200 years and CO₂ readings so that pupils can graph the results and draw conclusions."

"My teacher ask for next class to find the location of our school in degrees, minutes and seconds, but I really don't know how to get it. Please help me."

"How can I obtain recent global land clearing statistics?"

"I am looking for data on the Marianas Trench. Has it been changing in any way?"

"I'd like to know about atmospheric conditions (ex. Great Northern Lights) and be alerted as to when they are."

"Hi. My name is _____. I'm a senior and I'm working on a science project. My topic is the North Atlantic Oscillation and how it affects the weather in Europe. One of my friends has been talking to you and you sent her some very helpful data sets and I was hoping maybe you could do the same for me! If not it's OK, but anything you think you could do to help me would be greatly appreciated!! Thank you so much!"

Some typical user feedback:

Liam O'Reilly (Student) in search of rainforest destruction information for a report. "Thank you very much for the helpful information. I will write back to you and tell you what I get on it."

Professor Don Hyndman in search of a Weather Map and Jet Stream information.

"Thanks for the web links. The NOAA one is the best. Although I had browsed various NOAA sites I had not seen this one. It has some good images and explanations. One of them is actually a nice, concise explanation of the Jet Streams:

<http://www.srh.weather.gov/jetstream/global/jet.htm>"

Mariana Castillo in search of national map projection parameters for Bolivia. "It's amazing, thanks a lot!"

Arindam Ghosh in search of Climate data links for the Indian stations and the Indian Ocean. "Thanks a lot for the information. These links are really very useful. Thanks again"!

Charles T. Robinson had many questions pertaining to the GCMD parameter keywords. "Thanks for the quick response, this is exactly what I needed".

Carol Kreis (a freelance writer) developing an educational map for Newsweek about human impacts on climate change. She wanted to know how to obtain permission to reproduce the graph, "The increase of atmospheric CO2 over time". After receiving assistance, she wrote, "Thanks for your prompt reply".

Liesbeth Van Iseghem (Student) needed assistance on researching the Northwest Passage. She replied "It has been a while but you sent me information and you supported me in writing my term paper about the Northwest Passage. I want to thank you very much for your help. It was not an easy subject but it was so interesting. I learned so much and really felt a discoverer. At school there were people who supported me and were interested. There were, unfortunately, also 'non-believers', who thought it was a stupid subject. But I made it, yesterday I graduated with distinction!! I will always, from now on, follow a little bit the evolution in the Arctic sea-ice and

all related subjects! Thank you so much for your help and support, without you I could never succeed in doing this.”

Antje Dun (Australian Conservation Foundation Representative) requested assistance with land clearing statistics for 2002 and 2003. After receiving assistance, the user responded, “Thank you so much for your help. You are wonderful! I look forward to hearing from you if you find any more information about 2003.”

Mathew Koelz was in search of an algorithm for distance between a pair of lat/long coordinate. User responded, “Much appreciated... and thanks for taking the time to look... my calculator uses the cosine method, with a small addition for the radius for the average height amsl, but my problem was the distance for the minute of arc. With the information in the document, I hope to improve my calculator accuracy somewhat, which is only used for distances of 10000' or less.”

Helen Pan was in search of satellite and space photographs of China. She responded, “Thanks a million. It is powerful. I can't get these pictures in China.”

Hailie Cook was in search of global warming information. The user replied, “Thank you so much for the websites you sent.”

Mallory Guerin was in search of UV-B radiation and rainforest information. The user responded, “Thank you so much Gene for taking time out to help me. I think that I will be able to use this information in my paper.”

Danna Smith was in search of North Atlantic Oscillation data and responded, “Wow!!! Thank you so much!! This is a huge help! You have no idea how much I appreciate this!! Once again thank you sooooo much!!”

VIII. Collaborations

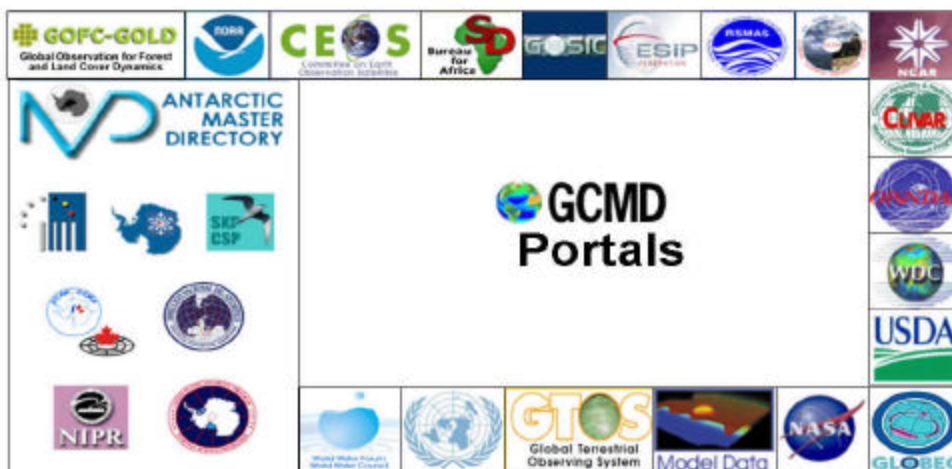
A. Portals

Many organizations acknowledge the importance of metadata related to their area of interest but do not have the resources required to manage their content. Portals allow organizations to maintain and document their data without duplicating the effort in creating another online directory. By hosting metadata through a GCMD portal, users can view the virtual subset of interest. Portals help provide science, or application-specific foci for other agencies, science focus groups, consortia, etc. and may be trademarked with the logo of an organization.

Using a portal to search a virtual subset is advantageous in that as metadata is added to the subset, it is also freely available from the GCMD's general search pages for scientists in other disciplines to access and use. Maintenance and creation of customized free-text and keyword search portals for the science community continued in 2003 - with eleven new portals added. Usage statistics are also regularly tracked and available.

See: http://gcmd.nasa.gov/Data/portal_index.html

GCMD Access to Community-focused Data



Maintenance and creation of customized free-text and keyword search portals for the science community continued in 2003. Additional portals for the Antarctic Master Directory, the Japan Aerospace eXploration Agency (JAXA), and other organizations will be added during 2004.

See: http://gcmd.nasa.gov/Data/portal_index.html

Current portals include:

- U.S. Department of Agriculture *created October 1999*
- Global Observation System Information Center (GOSIC) including Global Climate, Terrestrial, and Ocean Observing Systems (GCOS, GTOS and GOOS) *created March 2000*
- Joint Committee on Antarctic Data Management/Antarctic Master Directory (JCADM/AMD) *created August 2000*
- Distributed Oceanographic Data System (DODS) *created August 2000*
- Global Observation of Forest Cover (GOFC) [a CEOS project] *created September 2000*
- Earth Science Information Partnership (ESIP) *created December 2000*
- Global Ocean Ecosystems Dynamics (GLOBEC) [an IGBP project] *created May 2001*
- Committee on Earth Observation Satellites (CEOS) *created June 2001*
- World Data Centers (WDC) *created June 2001*
- Rosenstiel School of Marine and Atmospheric Science (RSMAS) [University of Miami] *created July 2001*
- Climate Variability and Predictability (CLIVAR) [a WCRP project] *created January 2002*
- Geographic Information for Sustainable Development (GISD) [a CEOS project] *created June 2002*
- Model Output Data portal *created September 2002*
- National Oceanic and Atmospheric Administration (NOAA) *created July 2002*

Created in 2003:

- Argentina Antarctic Center *created August 2003*
- Belgian Antarctic Program *created January 2003*
- Canadian Polar Commission/Canadian Committee for Antarctic Research *created October 2003*
- Finnish Antarctic Program *created January 2003*
- Swiss Committee on Polar Research *created November 2003*
- NASA Geospatial Framework *created September 2003*
- NASA GSFC GES Distributed Active Archive Center *created June 2003*
- National Center for Atmospheric Research (NCAR) *created September 2003*
- Remote Sensing for Conservation *created May 2003*
- United Nations (UN) Earth Science Data *created April 2003*
- World Water Forum (WWF) *created January 2003*

B. Collaboration Updates for 2003

DODS/OPeNDAP

The Distributed Oceanographic Data System/Open-source Project for a Network Data Access Protocol (DODS/OPeNDAP) currently promotes the GCMD through the portal on their website and through presentations at scientific conferences. The GCMD continues to serve as a locator for all of the DODS datasets. Statistics for this portal are presented to DODS each month. As of December 31, 2003, an average of 304 visitors accessed the portal each month. The GCMD has also worked closely with DODS to provide direct access to the Live Access Server (LAS) datasets and to provide a direct link through GCMD's OpenAPI. DODS recently introduced the OPeNDAP Data Connector (ODC), a stand-alone Java program, which allows the search and retrieval of datasets published by OPeNDAP data servers. The ODC allows one to find OPeNDAP related datasets and import them into client applications like IDL, Matlab, SPSS, Excel, or into databases such as Access and Oracle, and plot them with advanced graphics capabilities. The ODC Search Frame provides interfaces to the GCMD and the DODS dataset list. For additional information about the ODC, please see: <http://dodsdev.gso.uri.edu/ODC/>

USGS/Biological Resources Division



The following new metadata (182 records) were provided to the USGS/BRD National Biological Information Infrastructure (NBII):

- RAMAS Ecological software system - 3
- Biota of Virginia (BOVA) database - 10
- Biologically Integrated Geospatial Salmonoid data Access and Management (BIGSAM) - 13
- University of Tennessee - 16
- Patuxent Wildlife Research Center (PWRC) software - 9
- PWRC data - 4
- Brooklyn Botanic Garden Herbarium Catalog
- New York Metropolitan Plant Occurrence
- Patuxent Software metadata - 9
- Leetown Science Center - 19
- Southern Appalachian Information Node - 94
- Amphibian Research Monitoring Initiative - 3

ESIP/FIND

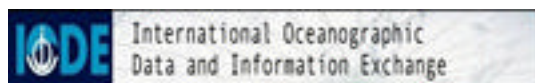


The Federation of Earth Science Information Partners (ESIP) brings together government agencies, universities, non-profit organizations, and businesses in an effort to make Earth Science information available to a broader community. Members of the Federation continue to provide DIFs and SERFs to the GCMD, which the GCMD also makes available through the Federation Information for Networked Discovery (FIND) on the Federation home page. The Federation added 685 new DIFs and 53 new SERFs for a total of 3817 DIFs and 156 SERFs in the GCMD. The GCMD is represented on the information and technology committee, e-journal committee, and the products and services committee. The GCMD staff members have also presented posters and given technical workshops at Federation meetings.

AMD/JCADM



The Antarctic Master Directory (AMD) is an effort coordinated with the Joint Committee Antarctic Data Management (JCADM) to offer Antarctic data collected by researchers from 15 of the Antarctic Treaty countries. Data may be searched through the AMD portal, which was created in 2001 and is located at: <http://gcmd.nasa.gov/Data/portals/amd/>. In 2003, portals were added for the following AMD nodes: Argentina, Belgium, Finland, Canada, Switzerland, Japan, and the United States. These portals are available at: http://gcmd.nasa.gov/Data/portal_index.html. Over 593 additional new DIFs were received from the following nodes: Australia, United States, Argentina, Spain, United Kingdom, Estonia, and New Zealand. In addition, the Australian Node (AADC) has installed the final MD8 software and tools.



IODE

Representatives from the GCMD coordinated with [International Oceanographic Data and Information Exchange](http://www.ioode.org/) (IODE) for the release of the MEDI metadata tool to the IODE community. The [MEDI tool](http://www.ioode.org/) is available online from the IODE website. GCMD continues to actively participate in IODE technical and scientific meetings. Meeting notes and proceedings are available online at <http://www.ioode.org/>. As a part of the outreach effort of the IODE community, the GCMD is currently featured within the Ocean Portal (<http://www.oceanportal.org>), a high-

level directory of ocean data and information related web sites. The objective of the ocean portal is to assist scientists and other ocean experts in locating data and information.



NOAA

Collaborations continued with NOAA during 2003. The entire National Climatic Data Center metadata content was retrieved from the NOAA Server, translated from FGDC/CSDGM to DIF, and loaded into the GCMD (completed April 2003). An effort was also initiated to replace outdated NOAA metadata content from other NOAA centers with updated information. A NOAA portal (<http://gcmd.nasa.gov/Data/portals/noaa/>) was created to feature over 1,586 NOAA datasets referenced in the GCMD. The “model output” portal, created in 2002 as a result of collaborations with the NOAA NOMADS project, exceeded 600 metadata records. See: <http://gcmd.nasa.gov/Data/portals/models/>



NCAR

NCAR

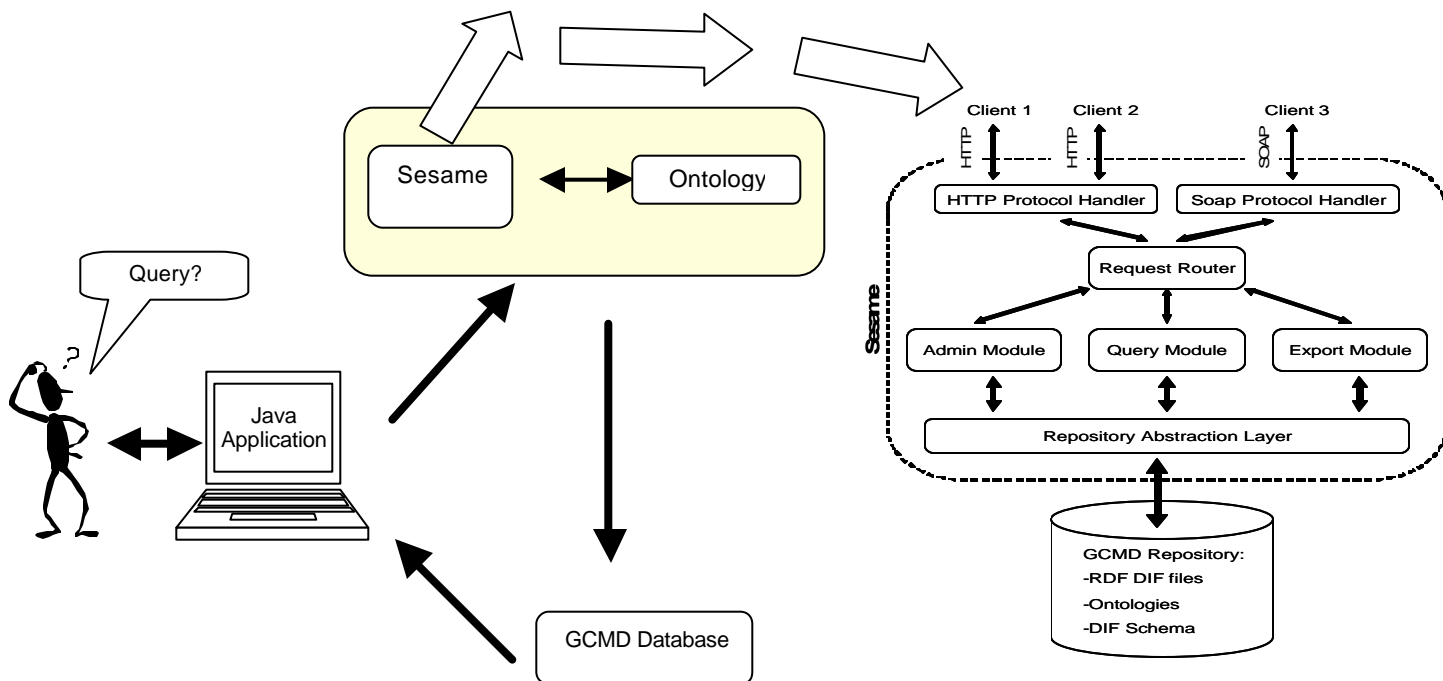
Collaborations continued with the National Center for Atmospheric Research (NCAR). A portal was created to feature data offered by NCAR <http://gcmd.nasa.gov/Data/portals/ncar/>. The GCMD staff also began working with NCAR’s Community Data Portal (CDP). The CDP staff created an XSL Style Sheet to convert their XML metadata to DIF XML. As a result, 34 new DIFs were contributed to the GCMD from CDP. The total NCAR metadata collection in the GCMD reached 472 DIFs in 2003. Additional DIFs are expected in 2004, as NCAR’s Data Support Section (DSS) is expected to archive almost the entire archive of European Centre for Medium-Range Weather Forecast (ECMWF) ERA-40 Model Data.



GLOBEC

Collaborations with the [Global Ocean Ecosystems Dynamic Program](#) (GLOBEC) community have continued. GLOBEC has continued to contribute to the [GLOBEC portal](#) by creating records from data collected in the Georges Bank and the Southern Ocean. The project has continued to promote their portal by featuring the GCMD within their website (GLOBEC Metadata Inventory and Information: <http://www.pml.ac.uk/globec/Data/metadata.htm>).

Graduate Summer Student Program



Through the work of a student in the Graduate Summer Student Program (GSSP) at Goddard Space Flight Center, the GCMD experimented with database structures in machine parsable format, which are associated with the semantic web. The semantic web provides the ability to search for semantic relationships among any DIF terms within an ontology, without the need to change the database structure when new classes and relationships are added. The real advantages occur when an ontology is enriched. The diagram above illustrates “Sesame” (an open source: RDF schema-based repository and query facility) middleware and its flexibility as a GUI or API – using alternative databases.



Mr. Viktor Pusztai from the United Nations Environmental Programme/Global Resource Information Database (UNEP/GRID) – Budapest (Hungary) has been working to integrate the PostgreSQL database into the MD software. His valuable contribution through the Committee on Earth Observation System International Directory Network (CEOS IDN) community will provide an open-source alternative to the Oracle database. Mr. Pusztai has taken the current SQL scripts to create the tables, indexes, views, triggers, etc. and is rewriting them in order to set up the same database structure in PostgreSQL. He is also making modifications to the MD software necessary to interact with this database.

University of Virginia

The University of Virginia group, led by Dr. Worthy Martin, presented the final report on their thesaurus work as part of an Earth Science Technology Office (ESTO) award. The thesaurus is currently running on a development machine at GCMD. It has the ability to query the Oracle database for terms and allows users to edit the terms already in the database. However, the thesaurus is not able to accept new terms to be added to the database.

GOSIC/GTOS

The Global Observing System Information Center (GOSIC) at the University of Delaware, Lewes coordinates the data activities of the Global Observing System (GOS), which consists of the Global Terrestrial Observing System (GTOS), the Global Ocean Observing System (GOOS), and the Global Climate Observing System (GCOS). The GCMD, in collaboration with GOSIC and GTOS representatives, identified new science keywords and added them to the GCMD database. Scripts were created to allow the direct query of DIFs from GOSIC interfaces. In October 2003, the GCMD presented a status report on the GCMD-GOSIC collaboration to the GOSIC Science User Working Group.

IOOS/DMAC

The [Data Management and Communications](#) (DMAC) subsystem will combine all of the global and coastal components of the sustained Integrated Ocean Observing System (IOOS), and link every part of the observing system from the instruments to the users. Representatives from the GCMD, in collaboration with the IOOS/DMAC Metadata Working Group, contributed to the metadata section and reviewed other sections of the IOOS/ DMAC Implementation Plan.

U.S. Office of Management and Budget

As requested by the U.S. Office of Management and Budget (OMB), the GCMD developed a NASA Geospatial framework portal - highlighting geospatial data investments consisting of framework data layers. These framework layers include cadastral, digital ortho imagery, elevation bathymetric, elevation terrestrial, government units, transportation, hydrography, and geodetic controls. See http://gcmd.gsfc.nasa.gov/Data/portals/nasa_geo/.

IX. Standards

A. GCMD Controlled Keywords. Although it is unclear how many groups use the GCMD keywords, 12 additional organizations that use them were identified in 2003 - bringing the known total to 23. The GCMD's Earth science parameter and services controlled keywords are available at <http://gcmd.gsfc.nasa.gov/Resources/valids/> and are currently used by:

- CEOS Interoperability Protocol (CIP)
- EOSDIS Data Gateway (EDG)
- EOSDIS Core System (ECS)
- NASA's Visible Earth
- NASA Taxonomy
- Federal Geographic Data Clearinghouse (FGDC)
- CCRS/GeoConnections
- Mercury/Beja Flor (ORNL)
- NASA Visible Earth/Earth Observatory
- Distributed Oceanographic Data System (DODS) (also known as OPeNDAP)
- NOAA (used in NOAA FGDC records as the thesauri keywords)
- Mermaid (formally CORIS)
- U. California Natural Reserve System
- Semantic Web (SWEET at JPL)
- DLESE (not explicitly using them but acknowledges GCMD as a resource when constructing high-level DLESE controlled vocabulary)
- AGU Cryosphere keywords (Larry Hinzman)
- GCDIS (Global Change Data and Information System)
- GOSIC (Global Observing System Information System)
- MEDI (Marine Environmental Data Inventory)
- BENEFIT (this is the Marine Fisheries system from South Africa)
- Neptune (Australia National Oceans Office) (<http://neptune.oceans.gov.au/>)
- SAIRE at NASA/ARC (<http://saire.arc.nasa.gov/GCMD-Final/key.html>)
- GCMD vocabulary considered during MBARI metadata workshop, October 2003.

During the year, the following additions and modifications were made:

New/Modified Science Keywords:

- Added new TOPIC: Climate Indicators
- Added 23 new Variables and 7 new Terms
- Currently (2003): 13 Topics, 117 Terms and 1219 science Variables

New/Modified Services Keywords:

- Added 1 new Variable
- Currently (2003): 8 TOPICS, 43 TERMS and 57 Services Variables

B. Data Set Content Description Standard - The Directory Interchange Format (DIF):
http://gcmd.nasa.gov/User/difguide/write-a-dif_v8.html

C. Analytical Resources Content Description Standard - The Services Entry Resources Format (SERF): http://gcmd.nasa.gov/Services/write_a_serf.htm

D. ISO 9001 for configuration control: http://gcmd.nasa.gov/Aboutus/software_docs/config.html

E. ISO Geospatial Metadata Standard - 19115/TC 211

GCMD representatives monitored and responded through the international community during the ISO standard process. In preparation for compliance with this standard, a cross-mapping was completed between DIF and ISO19115/TC211, and additions will be made to the DIF for compatibility. Mappings were also made between the ISO Topic and GCMD keywords. Proposals were sent to the CEOS Interoperability Forum, documenting the changes necessary for ISO compatibility.

F. Federal Geographic Data Committee (FGDC) and the new Geospatial One Stop

The explicit decision to retain the DIF format for internal use, while making all DIF metadata available as FGDC compliant records, has proven to be an excellent one. It has provided the distinct advantage of preserving the validation of metadata through the database – a property not available through the Clearinghouse. It has also allowed NASA to maintain fields critical for the GCMD that are not available in FGDC. These fields are considered to be absolutely essential to the NASA mission. They help prevent duplicative entries and assist in the discovery of data sets and include: (1) Entry ID (unique identifier for every data set); (2) Entry Title; (3) Parameters (science keywords essential for "normalization" of the database); (4) Sensor (Instrument); (5) Source (Platform - e.g., a satellite); (6) Minimum/Maximum Altitude and Depth; (7) Temporal Resolution; (8) Project; (9) Data Set Language; (10) Originating Center; (11) Data Center Name (variant); (12) Data Center URL; (13) Multimedia Sample URL; (14) Multimedia Caption; (15) Related URL; (16) IDN Node.

The GCMD's controlled keyword vocabulary may be selected for use by the FGDC. Because FGDC has no formal keyword vocabulary, the keywords may be used as thematic keywords in Section 1.6.1.1, where the Theme Keyword Thesaurus allows participants to specify a keyword framework.

NASA has met the requirements for the Geospatial One-Stop Initiative by providing information on geospatial acquisitions. See:

<http://gcmd.gsfc.nasa.gov/md/lucene/luceneSearch.html>

G. ISO Geospatial Metadata Standard 19115/TC 211

The release of MD9-ISO in May 2004 will bring the DIF into compliance with this ISO standard.

H. National Biological Information Infrastructure (NBII)

Work continues through a formal, interagency agreement with the USGS Biological Resource Division (BRD).

I. Crosswalk or mapping with other standards:

<http://gcmd.nasa.gov/Aboutus/standards/>

- Dublin Core
- Federal Geographic Data Clearinghouse
- National Biological Information Infrastructure
- ISO 19115

J. Concurrent Versioning System (CVS)

The Concurrent Versioning System is used within the project for software configuration control.

K. The Zen of Object Publishing Enterprise (ZOPE)

ZOPE, written in Python, is used for internal document management and for the Committee on Earth Observation's (CEOS) International Directory Network (IDN) interactions.

X. Conferences, Publications, Invited Lectures in 2003

Ritz, Scott, February, 2003. "Locating and Referencing NASA Terra and Aqua Data and Data-Related Services Using NASA's Global Change Master Directory", Poster presentation at the 19th Conference on Interactive Information and Processing Systems (IIPS) session at the American Meteorological Society Annual Meeting, Long Beach, CA.

Olsen, Lola M., April 2003. "CEOS International Directory Network Newsletter", #15.
<http://gcmd.gsfc.nasa.gov/pipermail/interop/2003-April000011.html>

Olsen, Lola M., May 2003. "International Standards for Scientific Data and Open Source Products – An Analogy", presented at Science Data Centers Symposium, University of Maryland.

Stevens, T., June 2003. "Retrieving geospatial data and services using the NASA Global Change Master Directory", Poster presentation at the Towson University GIS Conference (TUGIS), Towson State University.

GCMD staff, June 2003. "GCMD Applicability For and Usage at the GES DAAC", presentation at the Goddard DAAC.

Northcutt, T. and L. Olsen, June 2003. "Protecting Goddard Network Resources Via Transparent Network Bridges and Port Knocking Techniques". Director's Discretionary Fund Presentation.

Olsen, Lola, July 2003. "Overview and Demonstration of the Global Change Master Directory. Existing and Emerging Capabilities of the Global Change Master Directory", ESSAC Subcommittee on Information Systems and Services (ESISS), Washington, D.C.

Reitsma, Femke and L. Olsen, August 2003. "Spatial Data Accessibility and the Semantic Web". Submitted to SCISW2003.

Olsen, Lola, August 2003. "CEOS International Directory Network Newsletter", #16.
<http://gcmd.gsfc.nasa.gov/pipermail/interop/2003-August/000013.html>

Koken Petra, Warren T. Piver, Frank Ye, Anne Elishauser, Lola M. Olsen, and Christopher J. Portier, 2003. "Temperature, Air Pollution, and Hospitalization for Cardiovascular Diseases Among Elderly People in Denver", Environmental Health Perspectives, 111:10, 1312-1317.

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, 2003. NASA/Global Change Master Directory, Earth Science Keywords, Version 4.2.2
http://gcmd.nasa.gov/Resources/valids/keyword_list.html