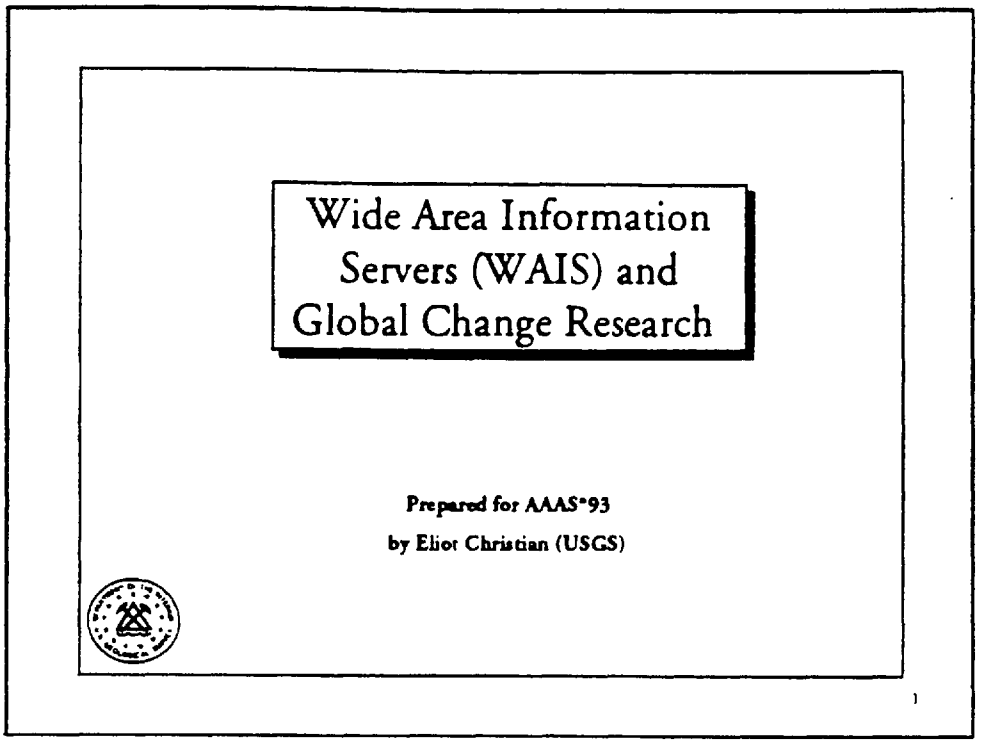


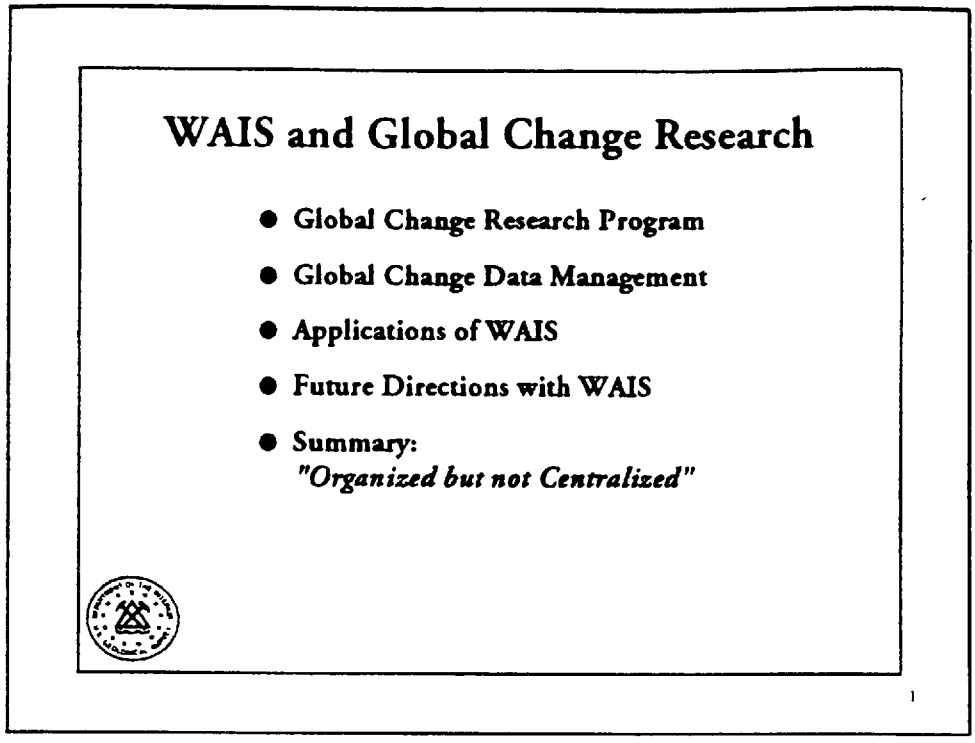
The goal of global change research is to document, understand, and predict changes in Earth systems on a global scale, providing a scientific basis for national and international policy formulation and decisions. Progress depends on the free flow of data and information, across national boundaries, scientific disciplines, and across many other intellectual and technical barriers.

For information to flow, a searcher must be able to discover where data and information exists and how it can be retrieved. However, researchers and information providers have traditionally worked within relatively closed communities, resulting in many separate and incompatible distribution mechanisms.

Fortunately, with the Z39.50 standard and adaptable implementations such as WAIS, we have a minimal cost approach that can greatly simplify the retrieval of data and information across divergent sources.

In this presentation, I'd like to give a sense of what the phrase "global change research" means, at least in terms of the U.S. program being led by the Federal government. For me, the diversity and incredible complexity of the program requires a simplified approach to information access that doesn't impose significant modifications on existing data and information systems.

It is my personal conviction that the Z39.50 standard and a WAIS-like approach must become an integral part of the global change research program.

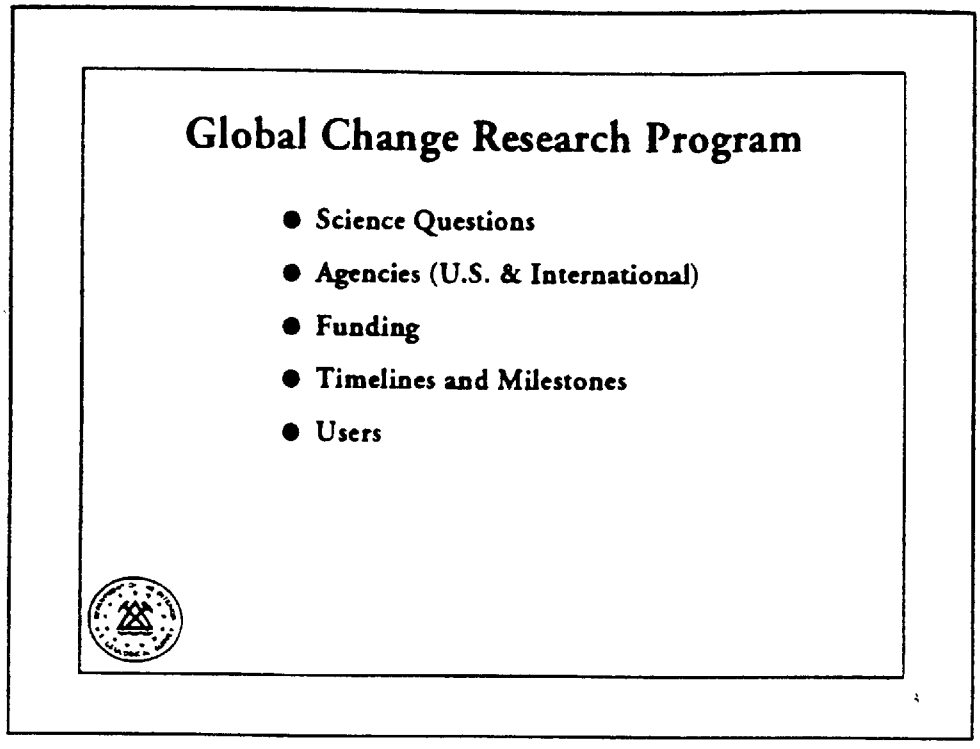


In outline, I'd like to present some background and ideas under five major headings. First, I'll discuss the overall Global Change Research Program, in terms of: the science questions being addressed; the major U.S. agencies participating in the program and roughly how the funding is distributed; the attempt to establish timelines and specific milestones; and my understanding of the user community the program should serve.

Next, I'll describe briefly the role of Data Management in the program. I'll draw out the distinction between "focused" and "contributing" programs, and introduce the "Global Change Data and Information Management Program Plan." This will lead into a discussion of the planned "Global Change Data and Information System."

I'd like to then describe my vision of how WAIS could be applied to the needs of the global change research program, especially in simplifying the discovery of data and information sources worldwide. I also see that there are aspects of the global change research requirements that call for further elaboration of information discovery and retrieval techniques, and I'll touch on some of these under the heading of "Future Directions with WAIS."

Hopefully, when I reach the Summary heading, you will understand why I've come to see the WAIS approach as essential to global change research. In a phrase, WAIS allows data and information to be "Organized but not Centralized."



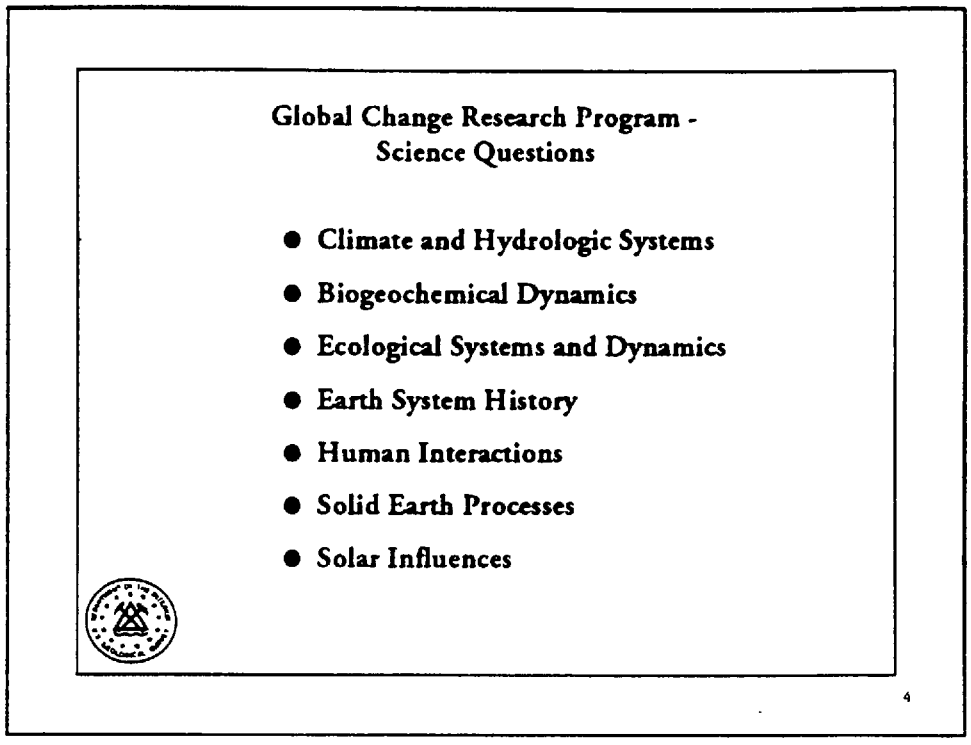
I'll be moving quickly through an overview of the Global Change Research Program. I'll start with a sample of some of the science questions being addressed, just to give a flavor of the range and complexity of the issues involved.

I'll then turn to which agencies are actively involved in the program, both within the U.S. and internationally. I think it's also useful to think about the many organizations, both governmental and non-governmental, that need to participate in the program on perhaps a less formal basis.

I'll present here just one breakdown of the U.S. program budget. While there are many different ways of cutting the funding pie to focus on different aspects, I've chosen a cut that highlights how data management fits within the program.

Under Timelines and Milestones, I'll sketch the current structure that tries to codify expectations of progress in the U.S. research program.

I'll conclude the program overview by coming back to the fundamental purpose of any program: to serve the needs of its users. Here, I'll describe the composition of that community of users, as I understand it.



The science questions being addressed by the U.S. global change research program are organized into seven areas: Climate and Hydrologic Systems, Biogeochemical Dynamics, Ecological Systems and Dynamics, Earth System History, Human Interactions, Solid Earth Processes, and Solar Influences.

Here's a sample of just a few of the dozens of broad science questions:

How will climate change affect temperature, precipitation, soil moisture patterns, and the general distribution of water on the land surface, and how will these changes in turn influence the atmosphere?

How do different coastal regions respond geologically and ecologically to rising sea level, and how can contributions from changes in climate be differentiated from those due to tectonic processes?

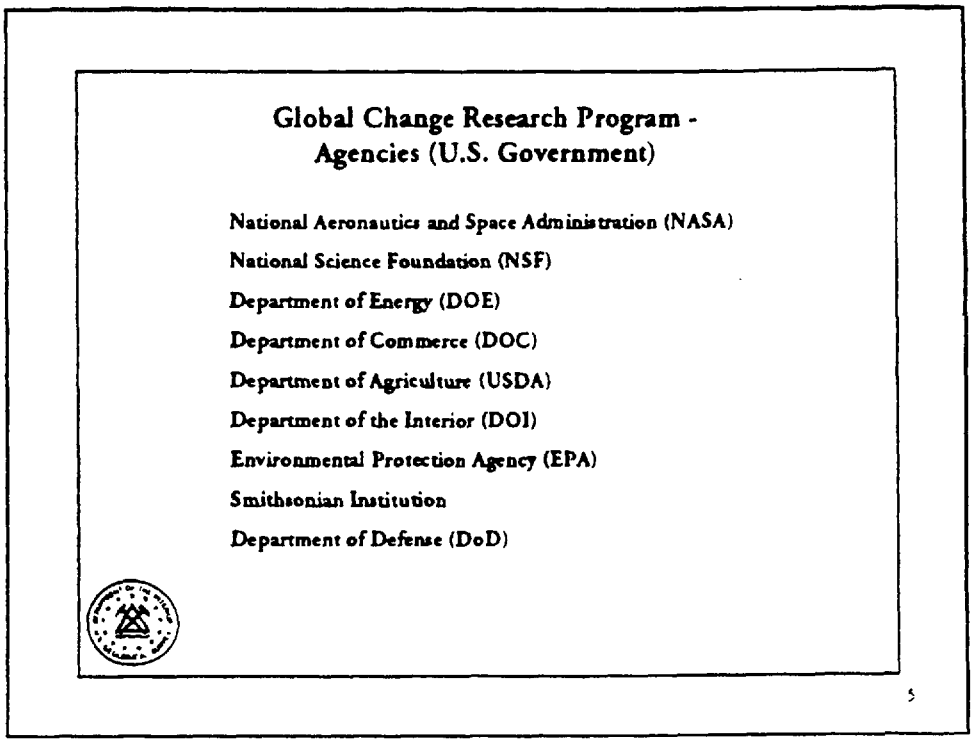
How do the oceans interact with the atmosphere in the storage, transport, and uptake of heat?

In which ecological systems and species are significant changes most likely to occur, and what attributes of importance to humans are at risk?

What are the natural ranges and rates of change in the climate and environmental systems?

What are the effects of volcanic emissions on regional and global climate?

What is needed to develop the information required to verify models of interaction between human and natural systems and to asses the likelihood of changes in those processes?



Here are the Federal agencies actively participating in the U.S. Global Change Research Program, listed in order of budget dollars:

National Aeronautics and Space Administration (NASA)

National Science Foundation (NSF)

Department of Energy (DOE)

Department of Commerce (DOC)

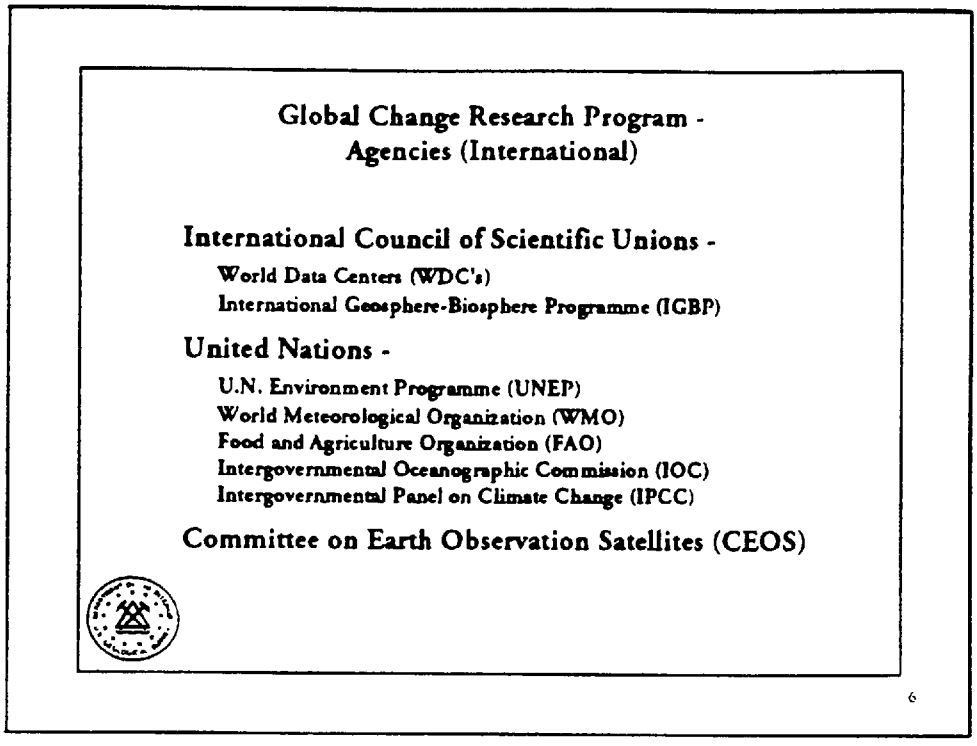
Department of Agriculture (USDA)

Department of the Interior (DOI)

Environmental Protection Agency (EPA)

Smithsonian Institution

Department of Defense (DoD)



Internationally, many different organizations are involved in global change research. Currently, the U.S. program tends to associate with formal organizations, such as the United Nations and the International Council of Scientific Unions (ICSU). Within ICSU, there is recognition that the system of World Data Centers is important to the success of the program. There is also a need to coordinate the U.S. program with the ICSU International Geosphere-Biosphere Programme (IGBP).

Among the many components of the United Nations that are important to the program, I've listed just a few: the U.N. Environment Programme (UNEP); the World Meteorological Organization (WMO); the Food and Agriculture Organization (FAO); the Intergovernmental Oceanographic Commission (IOC); and the Intergovernmental Panel on Climate Change (IPCC).

With the inherently international character of satellite remote sensing, and the importance of such systems to the research program, the U.S. program is also coordinated with the international Committee on Earth Observation Satellites (CEOS).

Missing from these lists of national and international organizations are thousands of other groups throughout the world that have various interests in, and contributions to, the overall global change research program. Consider, for example, the community of non-governmental organizations that convened at Rio de Janeiro during the Earth Summit. I believe that the collective impacts of these organizations will be as significant to long-term management of our home planet as will the governmental institutions.

Global Change Research Program - Funding ($ in millions)

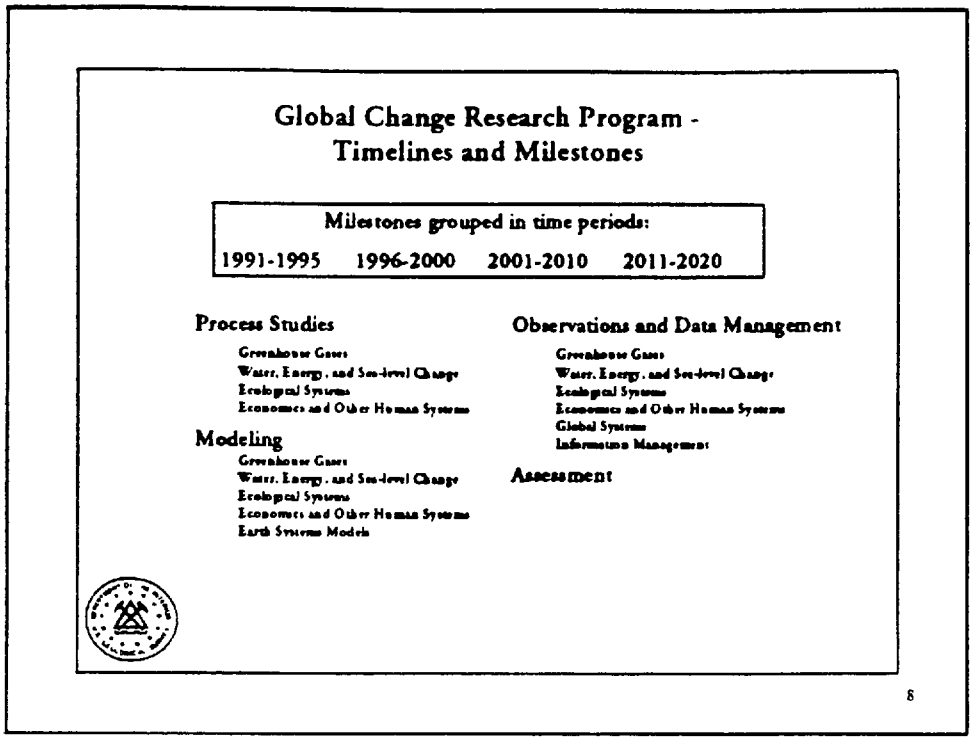
Science Elements	Actual 1991	Estimated 1992	Request 1993
Observation	259	254	329
Data Management	170	270	295
Understanding	457	468	606
Prediction	68	118	143
Total	954	1,110	1,372

Budget and other USGCRP documents available from:
Committee on Earth and Environmental Sciences
c/o National Science Foundation, Forms and Publications,
1800 G Street N.W., Washington, D.C. 20550

Under direction from the Committee on Earth and Environmental Sciences, agencies participating in the U.S. Global Change Research Program have presented their funding requests in various formats over the years. As of the 1993 budget request, one of the formats presented the science elements grouped under "Observations," "Data Management," "Understanding," and "Prediction."

As shown in this table, the U.S. program is at a level of about $1,400 million, with about $300 million ascribed to "Data Management."

If you're interested in further information on the U.S. program from a budget perspective, you can get copies of the series of reports titled "Our Changing Planet, the FY (1990, 1991, 1992, 1993) Research Plan." These reports are available from the Committee on Earth and Environmental Sciences, care of the National Science Foundation, Forms and Publications, 1800 G Street N.W., Washington, D.C. 20550.



In response to the Office of Management and Budget, the Committee on Earth and Environmental Sciences (CEES), has documented what the participating agencies expect to achieve in the research program. Given the huge investment, one would like some assurance that progress is being made, even though it is notoriously difficult to forecast where basic research will lead and what insights will be gained at what points in time.

Based on our current sense of which Earth systems might be affected by the activities of people, the program is now emphasizing the integrated study of:

sources and sinks of Greenhouse Gases and the processes that control them;

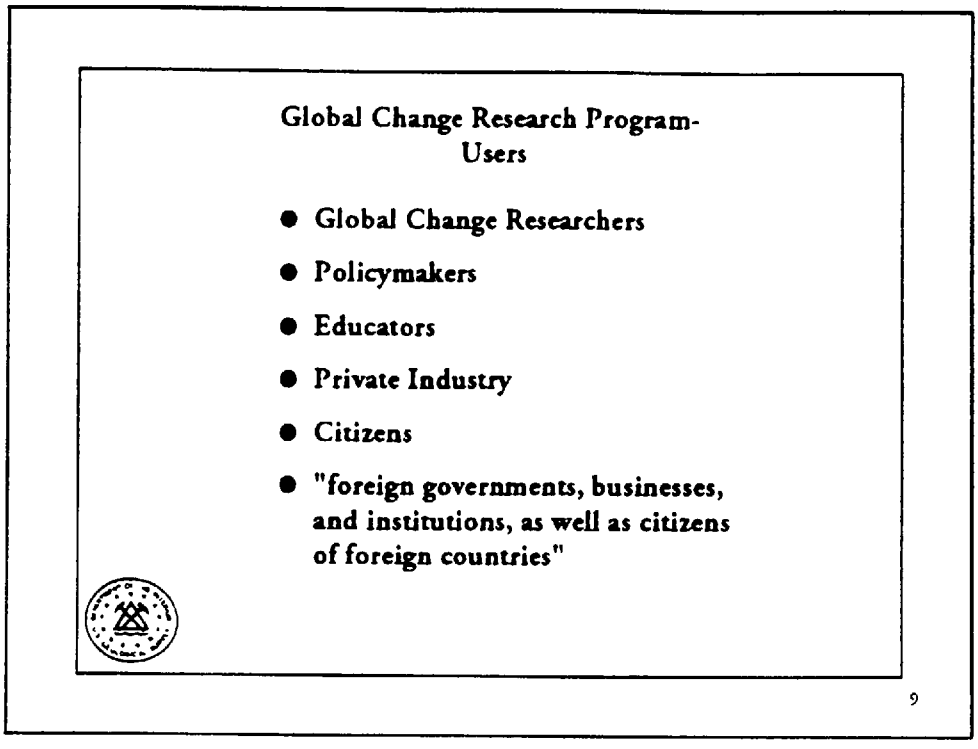
clouds, Water, Energy, and Sea-level Change, including radiative balance;

Ecological Systems, including species and ecosystem responses;

Economics and Human dimensions of global change, including resource economics, and the interactions and effects of human activities on global change; and,

global and Earth system observations and models to provide both global- and regional-scale data bases and models for predicting global change.

Milestones have been described for four time periods, reaching all the way out to the year 2020. They are arranged by integrating theme under Process Studies, Modeling, Observations and Data Management, and Assessment. Here, "assessment" refers to "assessing and synthesizing the state of scientific, technical, and economic knowledge and implications of global change to support national and international policymaking activities that cover the broad spectrum of global and regional environmental issues and to provide guidance for determining research priorities of the program."



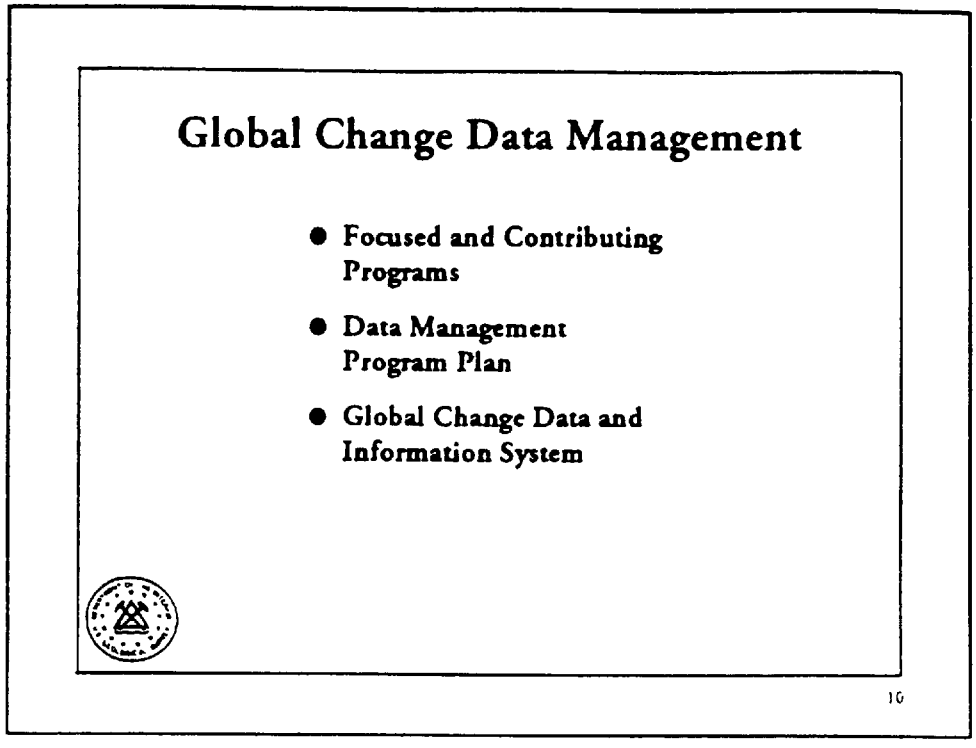
In the global change research program, the goal is not to discover new knowledge for its own sake--the program is directed to support national and international policymaking on a range of global and regional environmental issues.

Looking at the ambitious goals of the global change research program, I've drawn up a profile of the program's constituency--a list of users the program must serve.

Of course, the community of global change researchers must be well served to be full participants, and we've already talked about the emphasis on transferring knowledge to policymakers. Since the program will be underway for decades or longer, we must also involve the educational community if future generations are to build on today's work.

We already know that some aspects of modern consumer societies should be modified to lessen human impacts on Earth systems. To meet the myriad engineering challenges, and take advantage of emergent opportunities, we need to have the full involvement of private industry.

Ultimately, though, opinions and decisions of ordinary citizens decide the fate of the research program and any policies that flow from it. Also, the Global Change Research Act of 1990 specifically calls for the dissemination of U.S. global change research information, and information on technologies that mitigate or adapt to environmental changes, to "foreign governments, businesses, and institutions, as well as citizens of foreign countries". The focal point for foreign information dissemination is the Global Change Research Information Office now being established by the CEES.

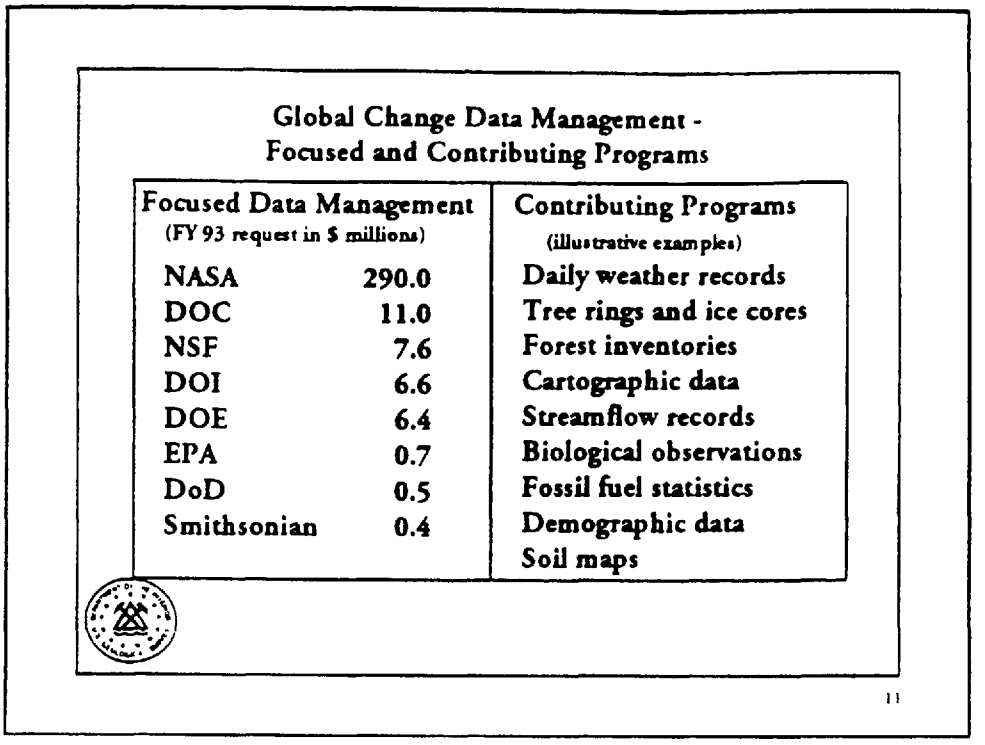


At this point, I'd like to move from an overview of the Global Change Research Program to Global Change Data Management, covering the specifics of data and information management in support of the program.

I'll start by drawing the distinction between "focused" and "contributing" programs, as these terms are used in the program. For me, this distinction is crucially important. It makes clear that the data and information component of the global change research program depends on lowering barriers not only among a few Federal agencies, but among a huge array of external organizations and systems.

I will then briefly describe the interagency "Global Change Data and Information Management Program Plan," to give a sense of the commitments agreed to by the Federal agencies participating in the program.

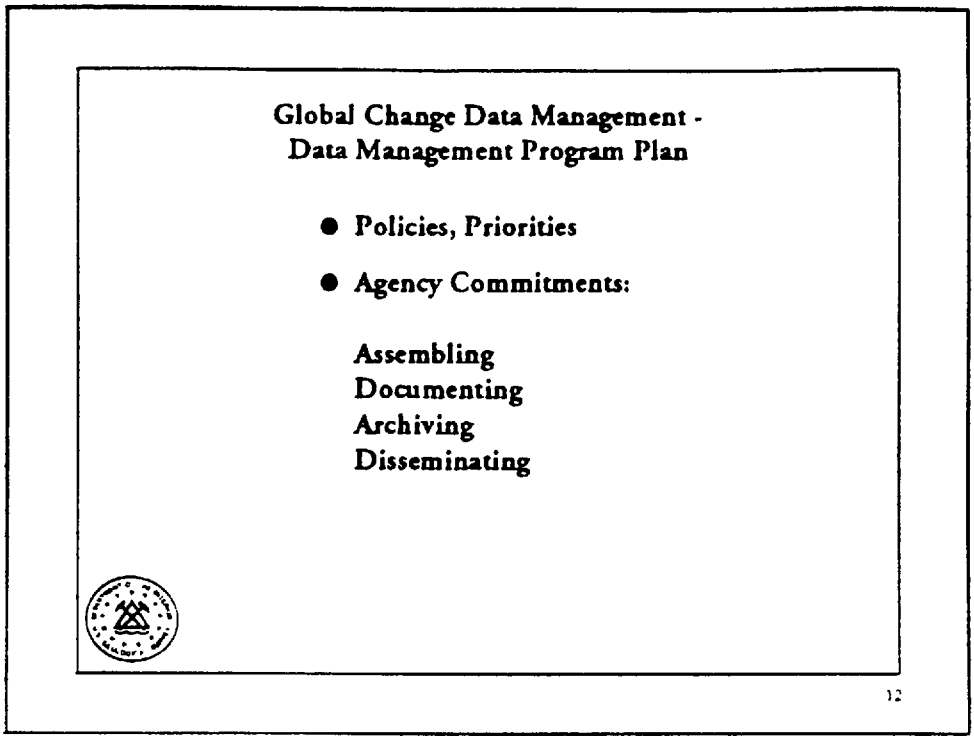
This will take us into a discussion of the interagency "Global Change Data and Information System." Since the implementation of this system is currently in the planning phase, I'll describe just a few of what I believe will be its salient characteristics.



As pointed out earlier, the Fiscal Year 1993 budget for the U.S. Global Change Research Program included a breakout of funding being requested for Data Management. The amounts by agency were as shown here: NASA $ 290 million, Commerce $11 million (NOAA is the largest component of this figure), National Science Foundation $7.6 million, the Department of the Interior $6.6 million (USGS is the largest component of the DOI figure), the Department of Energy $6.4 million, EPA $700,000, Defense $500,000, and Smithsonian $400,000.

These figures, though, only include the amounts actually allocated to perform data management under projects and programs directly funded by the U.S. Global Change Research Program. A much greater amount of funding is spent annually for a huge array of other projects and programs within the Federal government that are essential to the global change research program, although they are funded for other purposes. I've listed here just a few examples of the many sources of data and information crucial to global change research, yet not funded through the global change research program: Daily weather records; Tree rings and ice cores; Forest inventories; Cartographic data; Streamflow records; Biological observations; Fossil fuel statistics; Demographic data; and Soil maps. (We do not have good estimates of the actual funding for data management in contributing programs, but we do know it is larger than the $300 million allocated to focused data management.)

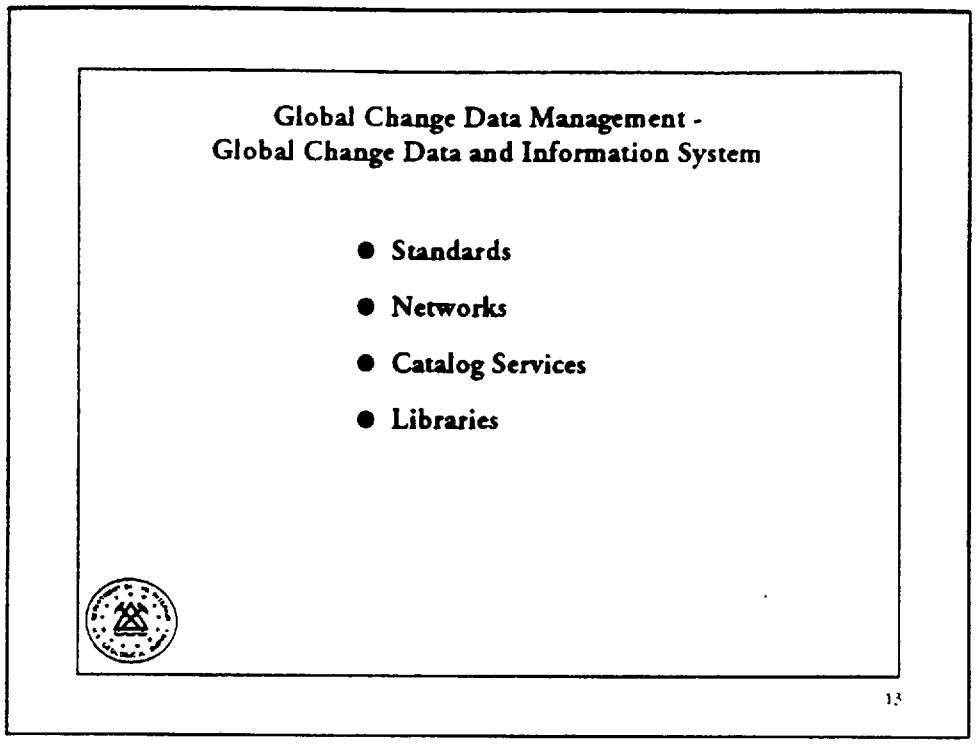
As you might imagine, it is a wholly different matter to achieve cooperation from organizations whose attentions are elsewhere than it is to coordinate those you fund.



Agencies participating in the U.S. Global Change Research Program are committed to an interagency global change data and information management program, consistent across agencies, and involving and supporting the university, international, and other user communities. The data and information management program is documented in a "U.S. Global Change Data and Information Management Program Plan." (Copies of this plan are also available from the CEES, care of the National Science Foundation.)

The Program Plan asserts the participating agencies' commitments to a common set of policies and a coordinated process for establishing priorities. It further commits the agencies to specific objectives in four activity areas comprising data and information management: assembling, documenting, archiving, and disseminating.

In assembling, there is a commitment to involve the research and user communities in identifying and assembling priority data sets and to seek cooperation with international partners to share global change data and information. Under documenting, the agencies are committed to publish sufficient documentation with data so that users are well informed about its quality. Agencies also commit to archive data and information for decades and centuries. And, in the area of disseminating, agencies are committed to maintain and facilitate various mechanisms so that it is timely, affordable, and easy for users to access and use global change data and information.

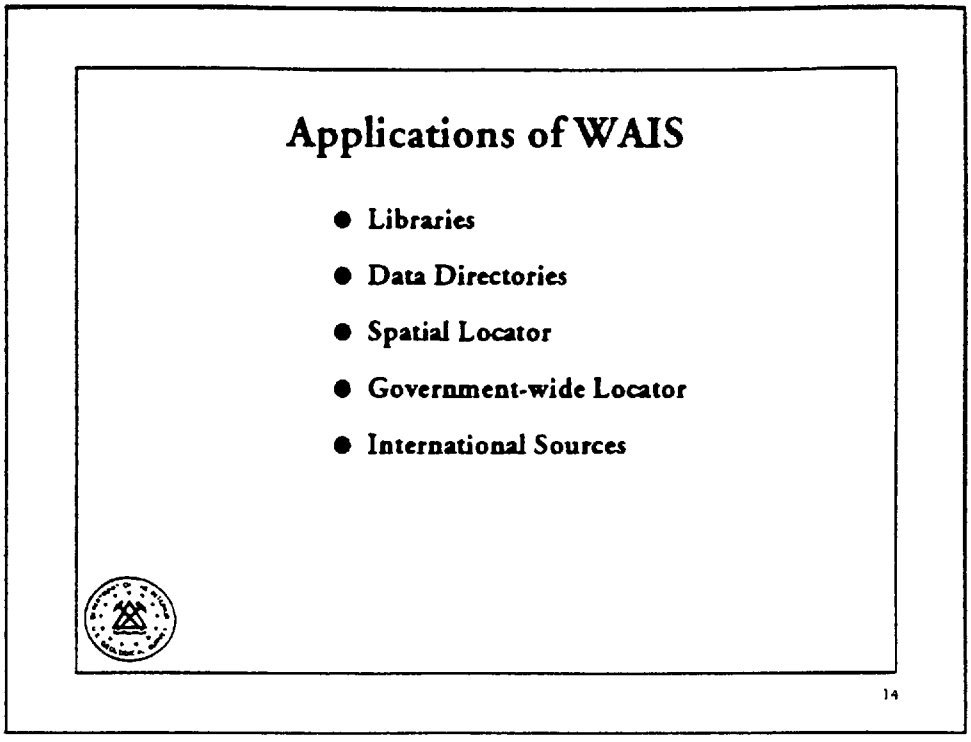


With the stated goal to make it as easy as possible for users to access and use global change data and information, participating Federal agencies are organizing a "Global Change Data and Information System." This is not some large, central computer center or data repository. Rather, it is distributed among existing facilities, with crucial standards agreed upon among the agencies. Each agency is expected to retain responsibility for its own mission data and information while providing certain services in concert with other agencies.

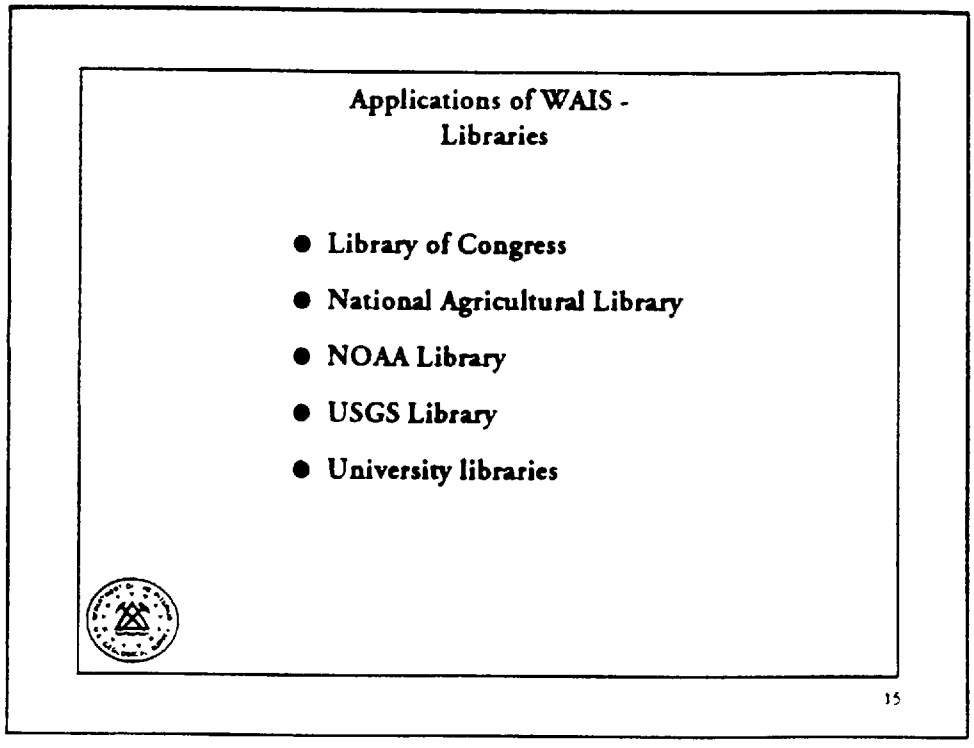
Existing and future networks will be a primary facility to provide researchers and other users with access to colleagues and to data and computation resources. No additional networks should be instituted for the Global Change Data and Information System, although there will be much attention to assuring that adequate interconnections exist and that navigating within and between the networks is made as simple as possible.

The main purpose of Catalog Services is to help users determine what data and information available throughout the world are useful for a specific task. This may require summary-level descriptions, points of contact, relevant publications, and sometimes more detailed information on appropriate uses and specific ordering assistance. Catalogs will be published on-line, off-line, and in paper formats.

Libraries are recognized as a basic resource for everyone needing information. A stated goal of the Global Change Data and Information System is that data and information formats, products, and network interfaces will be compatible with library systems.



Now, I'd like to describe my vision of how WAIS would be applied to the needs of the global change research program, especially in terms of simplifying the discovery of data and information sources worldwide. In this section, I'll touch briefly on the topics of libraries, data directories, the application of the WAIS spatial locator feature to global change research, the relevance of the Government-wide locator initiative, and tying into international sources of data and information.



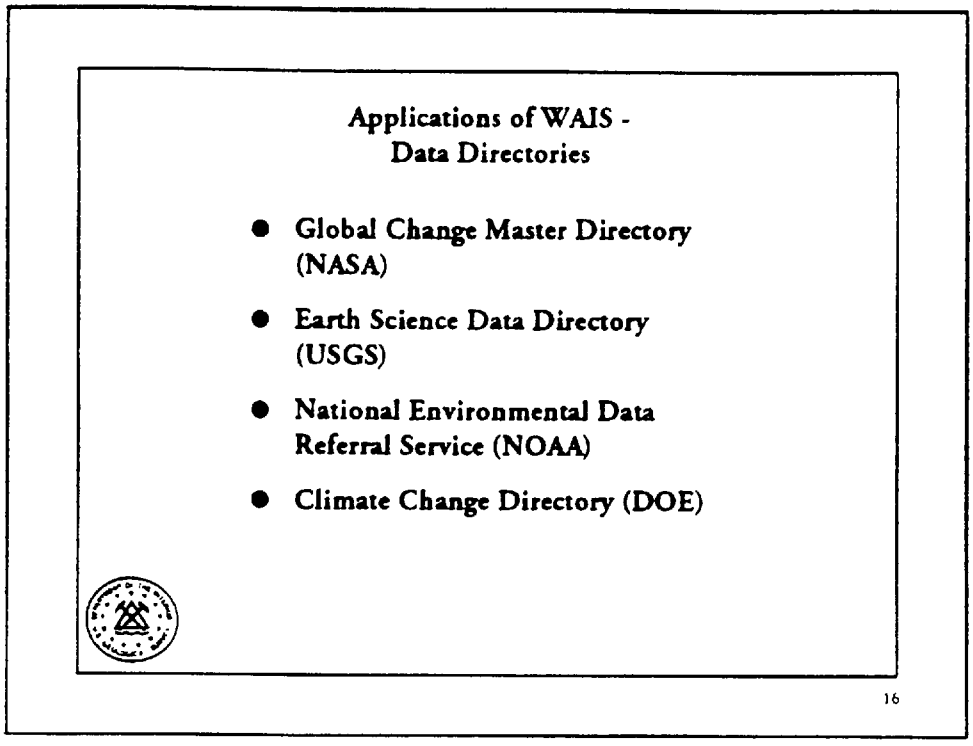
We must remember that libraries far outnumber data centers, and that libraries typically have few resources to invest in yet another automated interface. Fortunately, libraries already have a strong orientation toward using the Z.39 standards. For the needs of global change researchers, Z39.50 compatibility in catalog services is the only way I can see that we will be able to tap the vast resources of library information systems.

I've listed here just a few examples of U.S. libraries that the global change research user community will want to access:

the Library of Congress
the National Agricultural Library
the NOAA Library
the USGS Library
various University libraries

I think it is also very exciting to see that scientific journals are beginning to appear in an on-line electronic form. This will be a very important development for global change research, and we need to be positioned to take advantage of it.

In my opinion, the best way to prepare for electronic dissemination of scientific journals is to adopt the Z39.50 standard and help create appropriate attribute sets for these products, supporting formats such as SGML (Standard Generalized Markup Language).



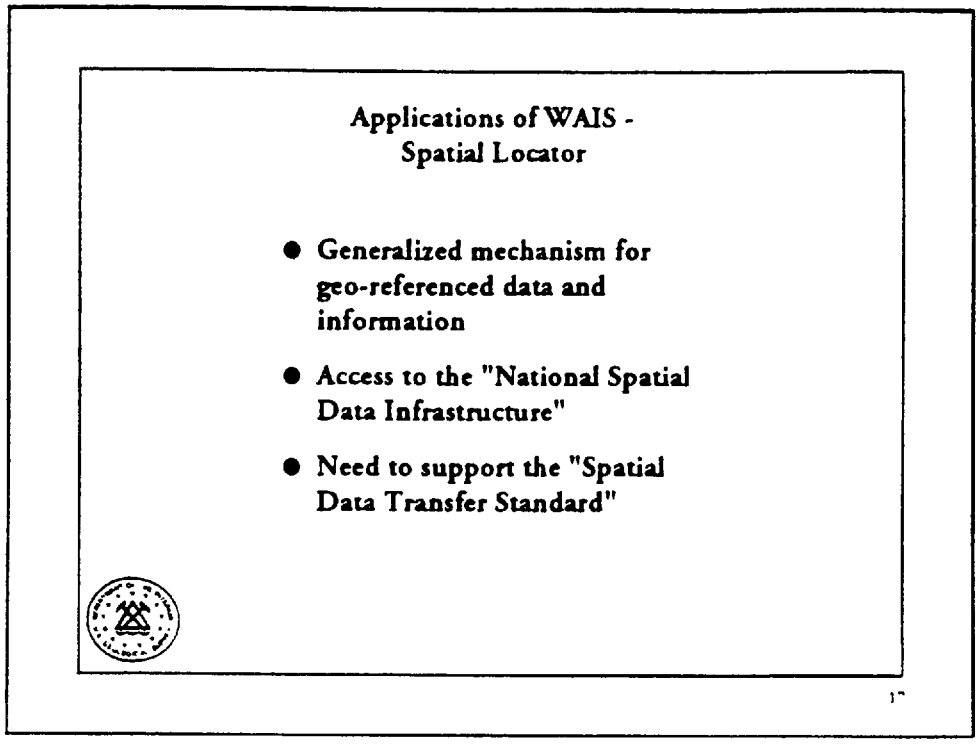
I am sure there exist many data directories that bear directly on the data needs of the global change research program. As a proof of concept, we have brought up as WAIS sources copies of some existing data directories pertinent to global change research:

the Global Change Master Directory, maintained by NASA
the Earth Science Data Directory, maintained by USGS
the National Environmental Data Referral Service, maintained by NOAA
the Climate Change Directory, maintained by the Department of Energy

Data directories are in a variety of formats, and most are maintained for purposes other than global change research. How can we encompass data and information directories maintained by the National Technical Information Service, the Patent and Trademark Office,the National Library of Medicine, and the Centers for Disease Control? Extending our vision to sources outside of the Federal government and then internationally as well, I think we'd find the same situation for hundreds of directories.

It seems to me, we cannot and should not attempt to consolidate all data directories. Nor should we organize them in some specific hierarchy that could only be appropriate for our current understanding of global change research needs. Rather, we should support the Z39.50 standard within the global change catalog services and help show managers of other data directories that it is in their interest to support Z39.50.

I think of directories as advertising--you don't try to reach everyone with just one ad. You present your product in different ways to different audiences. And, I find that most system managers jump at the chance to reach a bigger audience at minimal cost.

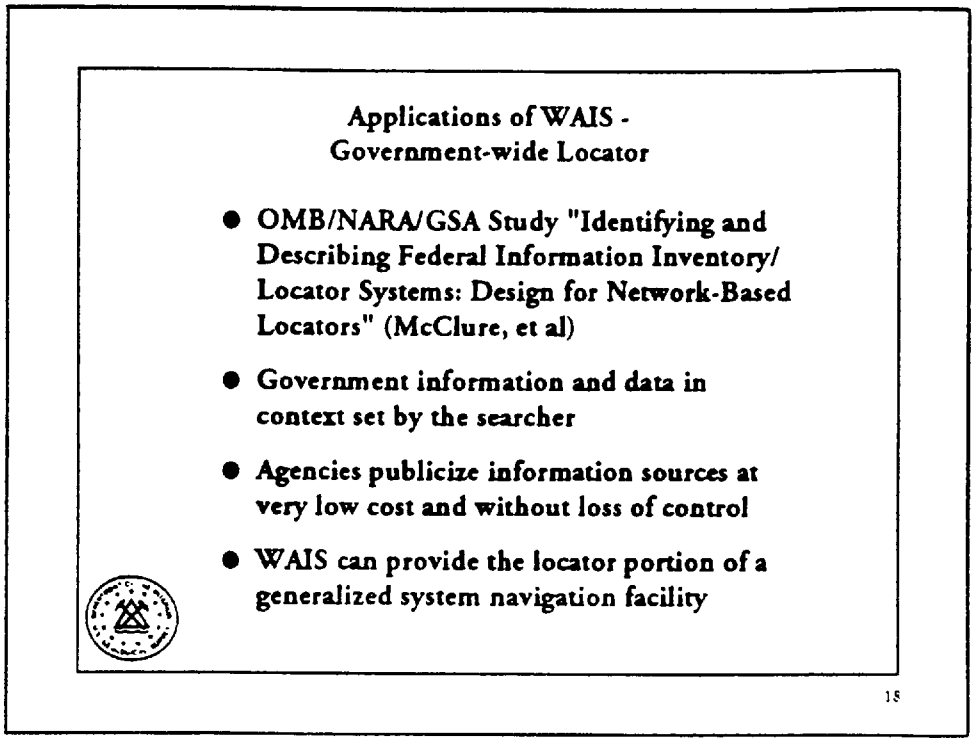


Much of global change research is concerned with data referenced to the surface of the Earth. The spatial locator feature of WAIS will be a very useful extension to he Z39.50 standard as a generalized mechanism for finding geo-referenced data and information.

My impression is that Z39.50 and a WAIS-like approach are very likely to be adopted by the Federal Geographic Data Committee in its implementation of a "National Spatial Data Infrastructure." This massive coordination initiative at the Federal level will likely influence mechanisms adopted by state and local governments, and the governments of some other nations as well.

Consensus has already been reached on an exchange format for spatial data--a format known as the "Spatial Data Transfer Standard." Once appropriate attribute sets are created for this exchange standard, users should be able to retrieve multiple map layers from a wide variety of sources and readily combine them in new ways.

This ability should greatly expand our collective ability to discern all kinds of patterns in Earth systems and human interactions over time and space.

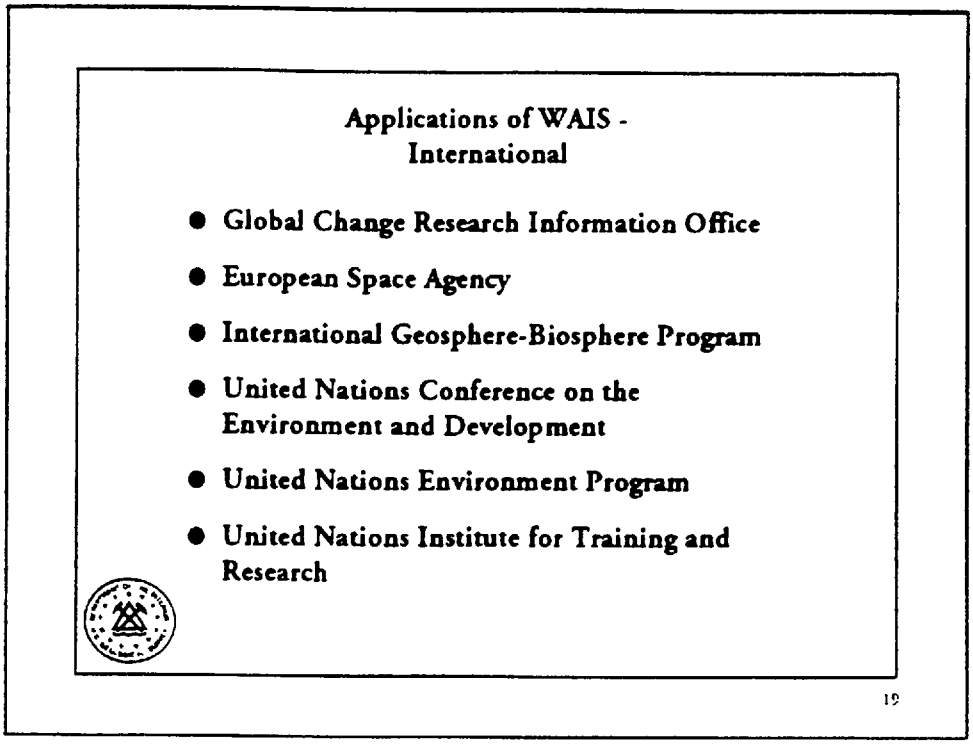


The Office of Management and Budget, the National Archives and Records Administration, and the General Services Administration have been concerned with increasing public access to government information and identifying existing Federal locator systems. In 1992, Syracuse University delivered a final report titled "Identifying and Describing Federal Information Inventory/ Locator Systems: Design for Network-Based Locators" by Dr. Charles McClure, and others. A key recommendation is that OMB should require agencies to have locators accessible over the Internet and conforming to the Z39.50 standard. Mechanisms to implement the recommendations are under consideration, but I believe Federal managers can anticipate strong encouragement to support the Z39.50 standard.

This new approach is far superior to the previous thinking that tried to establish a centralized catalog of all Federal data and information holdings. Not only is a central catalog perpetually out of date, its usefulness is inherently constrained to those uses envisioned by its designer. In contrast, the WAIS/Z39.50 approach allows the whole range of Government data and information to be viewed in a context set by the searcher.

The WAIS approach also allows agencies to publicize information sources at very low cost, and where the agency supports its own WAIS servers, with no loss of control.

It's important to remember that WAIS can usually be added onto existing systems without disrupting current processes. WAIS can also serve as a generalized system navigation facility, giving users a simple way to discover and access agency systems.



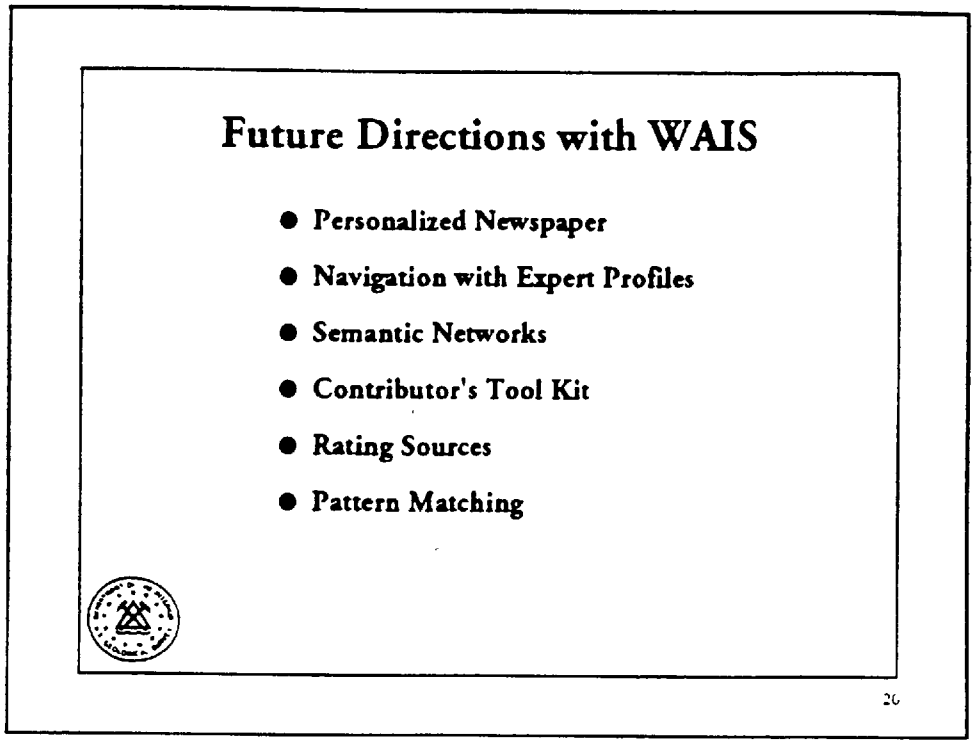
In the International arena, the problems of interfacing multiple data directories and foreign library holdings are greatly compounded by the lack of coordinating authority as well as by language barriers.

Here is just a very short list of some obvious international linkages that need to connect to the U.S. Global Change Research Program:

the Global Change Research Information Office
the European Space Agency
the International Geosphere/Biosphere Program
the United Nations Conference on the Environment and Development
the United Nations Environment Program
the United Nations Institute for Training and Research

Again, Z39.50 is very appropriate for the global change catalog services since it has an international standard counterpart. Although the WAIS implementation is based on the Internet, which connects over 10 million computers worldwide, it will also be important in some countries that access be supported over the Open Systems Interconnection communications protocols. The Z39.50 standard is fully compatible with those standards, as well.

We have heard for years that automation support for language translation will be coming soon. I don't know when we'll really have useful machine translation, but I am sure that market forces will make it available within the Z39 standards.

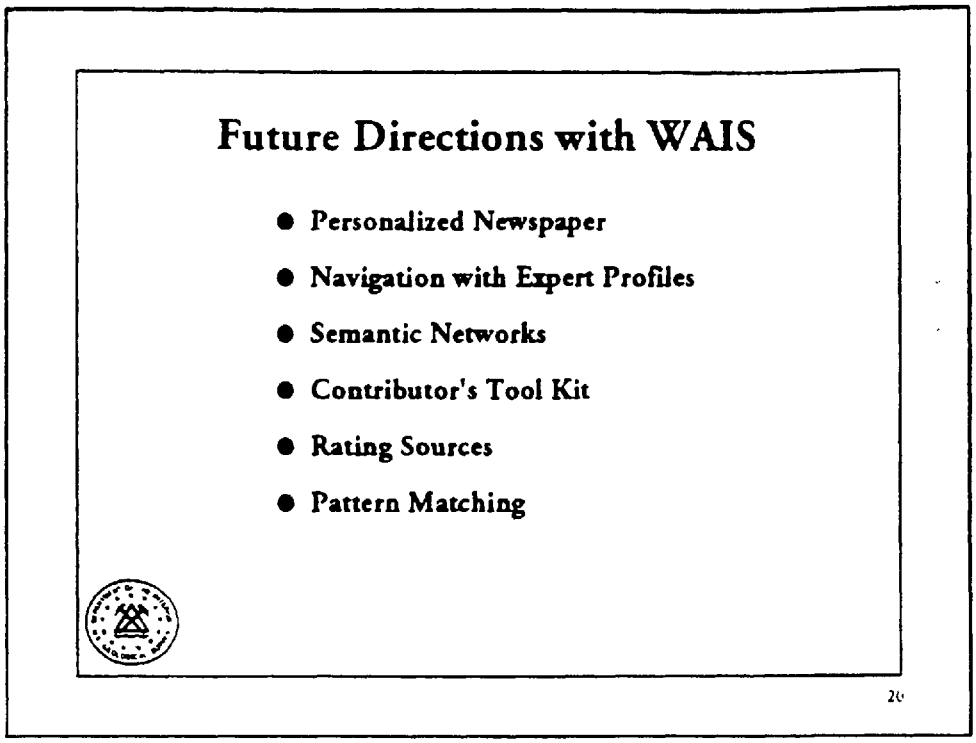


As I mentioned earlier, I think there are aspects of global change research that require further development of information discovery and retrieval techniques. For example, WAIS needs further elaboration of techniques for dealing with sources at all levels of aggregation. This is the same "ants and elephants" quandary that has dogged efforts to organize the global change holdings. The key to success will be to allow for the user to see things in an order that seems natural in his or her context, without overly constraining the ability of other users to see the same sources in a very different context.

Since we all expect the information explosion to continue, there is a critical need for better automated filters. One promising direction is automatic, unattended retrieval of information based on a profile of the user's interests. Such facilities are now becoming available under the general catch-phrase of "personalized newspapers."

An extension of this idea is that client software can monitor how the user actually handles information and develop a characterization that becomes part of a rule set that assists in discovering new sources of interest. As such "expert profiles" are developed, they could be combined or passed on to novice searchers. In effect, the expertise of a proficient information navigator becomes a valuable information product itself.

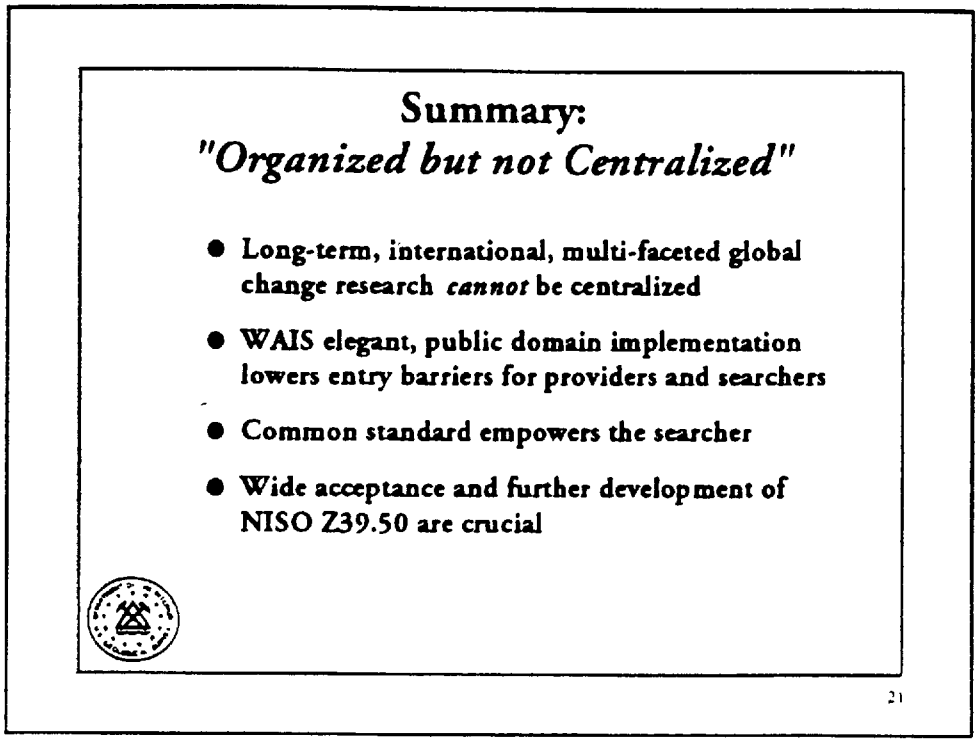
Concept searching and dynamic hypertext via semantic networks promise to greatly improve the accuracy of text searches, and might provide a path for overcoming language and jargon barriers. We will soon see these techniques in common use on Z39.50 clients and servers.



With the WAIS/Z39.50 approach, virtually anyone can become a publisher of data and information. Yet, we expect archival responsibilities to remain centralized. Data centers will continue to need contributors to provide documentation on data and information placed into the archive. We envision a "contributor's tool kit" on a WAIS client that would prompt for this documentation. This feature could be a button on the WAIS Query screen that opens up a fill-in-the-blanks form, does a spell-check, and either picks up the data online or presents directions on where to send it. The tool kit could also suggest an electronic peer-review for the documentation based on associations with the documentation of other holdings.

As information sources proliferate, there is a need for critical review of sources to guide users as to which sources are most appropriate for which uses. I envision that various organizations will establish sources acting as directories of other sources. The listing organization would include only those sources it considers worthwhile, and might also annotate the source description with an editorial on its strengths. A novice user might limit initial searches to sources listed in those directories run by organization with a good reputation in the field of interest (such as the U.S. Geological Survey for geology or NASA for space science).

At present, the searching part of the Z39.50 standard is oriented toward language patterns. I am hopeful that the standard will evolve toward more general pattern matching, to deal with images, time-sequenced maps, and mathematical patterns. I also expect servers to be used for distributing software, in effect "teaching" clients new tricks.



By now, I hope you share my sense that the Z39.50 standard and the WAIS approach could be very beneficial to global change research. In "sound bite" form, WAIS allows data and information to be "Organized but not Centralized."

To me, the fact that global change research is long-term, international, and multi-faceted dictates that the data and information systems supporting it must be de-centralized and based on international, very flexible standards for information search and retrieval. In my opinion, Z39.50 is ideally suited to this purpose.

I am an advocate of the WAIS public domain implementation of Z39.50 since it drastically lowers the up-front costs for providers and searchers. Of course, not everyone is comfortable with using public domain software, and good commercial implementations of WAIS and Z39.50 are now arriving on the market.

Given that the global change user community is so broad, I am wary of imposing very much structure on the process for information discovery and retrieval. To me, the Z39.50 standard allows us to empower the searcher to view the holdings from many different perspectives. Even so, it does not prevent information providers from providing more structured views whenever more structure is appropriate for particular users.

Because I am convinced that wide acceptance and further development of NISO Z39.50 are crucial, I have become somewhat of a WAIS-evangelist. Let me close by inviting all of you to the USGS in Reston March 19 for the first "Special Interest Group on WAIS", SIG-WAIS. An announcement of the SIG-WAIS meeting is included in your handout.

SECOND INTERNATIONAL SYMPOSIUM: NATIONAL SECURITY & NATIONAL COMPETITIVENESS: OPEN SOURCE SOLUTIONS Proceedings, 1993 Volume II - Link Page

Previous	To Demonstrate the Need for Australia to Develop a Strategic Policy on Open Source Information (OS!) Which Capitalises on Emerging Trends
Next	2025, Fisher & Pavkel. and Geelong District Water Board.. What Do They Have in Common ? - or Business is War

Return to Electronic Index Page