The GEO Data Sharing Challenge: Putting Principles into Practice

Alliance for Earth Observations Briefing 19 March 2007

Dr. Robert S. Chen
Director and Senior Research Scientist, CIESIN
Manager, Socioeconomic Data and Applications Center (SEDAC)
Secretary General, Committee on Data for Science and Technology (CODATA), International Council for Science (ICSU)

Center for International Earth Science Information Network
The Earth Institute, Columbia University
Palisades, New York, USA

e-mail: bchen@ciesin.columbia.edu

World Data Center for Human Interactions in the Environment

Background on CODATA

- CODATA: an ICSU interdisciplinary body
  - Established 40 years ago
  - Headquarters in Paris, co-located with ICSU Secretariat
  - More than 20 member countries, 15 scientific unions and interdisciplinary bodies
  - Long-standing involvement in data policy issues across the sciences
- CODATA committed to open access to data in key application areas such as disaster management, e.g.:
  - "The scientific community needs to press governments not only to release specific data sets that are vital to disaster management and planning, but also to establish a "good Samaritan" principle for the use of data and information in humanitarian emergencies."

EDITORIAL

Science and the Digital Divide

At the launch of the World Summit on the Information Society (WSIS) in Geneva in December 2003, the world community strongly affirmed the central role of science in developing an information society and affirmed the principle of "universal access with equal opportunities for all scientific knowledge and the creation and dissemination of scientific and technical information." The WSIS Declaration of Principles recognized the essential role of the public domain and public institutions such as libraries, archives, and museums in supporting the growth of the Information Society and providing free and equitable access to information.* The WSIS Plan of Action suggested numerous approaches to implement these principles, including "e-science" as a key application of information and communication technologies in support of sustainable development.†

The international scientific community succeeded in raising these issues at WSIS and securing widespread support from participating governments. Now, with the second phase of WSIS taking place in Tunis in November 2005, the scientific community needs to take the lead in demonstrating how science—and universal access to scientific data, information, and knowledge—can make a critical difference in sustainable development and overcoming the "digital divide."

The deadly South Asian tsunami in December 2004 and what many have called the "silent tsunamis" of millions of unnecessary deaths and untold suffering from malnutrition, disease, and poverty remind us that science has far to go. Scientists must work not only to predict future hazards and develop new medicines and vaccines, but also to make scientific data and information much more accessible and useful for real-world decision-making. These disasters underscore the need to better understand how societies can best organize themselves to address pressing problems posed by limited resources, conflict, poor infrastructure, and inadequate skills and knowledge. Scientists, the original developers of information and communication technologies, often take for granted their ready access to data and information, software and hardware, and networks of colleagues. But for billions of people, even the most rudimentary access to life-saving scientific expertise and knowledge, such as an early warning or a new cropping method, is a major challenge.

How can the international scientific community help reduce the digital divide? Already, many scientists and scientific institutions are working to improve the reach and effectiveness of science through information and communication technologies. The International Council for Science (ICSU) and its Committee on Data for Science and Technology (CODATA) are collaborating with WSIS to collect and document such efforts (www.wsis-online.net/science/home_EN/). But more needs to be done.

Scientists can support distance education and training; improve the accessibility of information and communication technologies to disadvantaged, marginalized, and vulnerable groups; communicate technical knowledge to the general public; and establish digital libraries, data archives, and other mechanisms to increase access to scientific information. We urge the scientific community to come up with more creative ideas and outcomes. Noteworthy examples on this front include the efforts by the Massachusetts Institute of Technology to provide electronic access to its course materials (http://ocw.mit.edu/index.html) and by the Global Biodiversity Information Facility to make primary scientific biodiversity data openly available (www.gbif.org). The scientific community should also consider new approaches to open electronic access, such as the Science Commons (http://sciencecommons.org), that, among other things, address the complex issue of licensing structures.

Immediately after the South Asian tsunami, critical data on elevation, population location, administrative boundaries, and damage could not be shared because of intellectual property and national security constraints. Even now, the 30-meter-resolution data from the Shuttle Radar Topographic Mission (SRTM) flown by NASA in the year 2000 is not publicly available, although it could potentially provide the best available elevation information regarding most of the world's coasts. The pending decision by the U.S. National Geospatial-Intelligence Agency to prohibit public access to various aeronautical products would be another step in the wrong direction. The scientific community needs to press governments not only to release specific data sets that are vital to disaster management and planning, but also to establish a "good Samaritan" principle for the use of data and information in humanitarian emergencies.

Science helped to create the Information Society—it can now help extend that society to all.

Shuichi Iwata and Robert S. Chen

Shuichi Iwata (University of Tokyo) is president of ICSU's CODATA. Robert S. Chen (Columbia University) is secretary-general of CODATA. CODATA is based in Paris, France.

10.1126/science.1119500

*WSIS, Declaration of Principles (document WSIS-03/GENEVA/DOC/4-E, 12 December 2003). †WSIS, Plan of Action (document WSIS-03/GENEVA/DOC/5-E, 12 December 2003).

Published by AAAS

Editorial in *Science* by Iwata & Chen 21 October 2005

Agreed GEO Data Sharing Principles

- There will be **full and open exchange of data, metadata, and products** shared within GEOSS, *while recognizing relevant international instruments and national policies and legislation*.

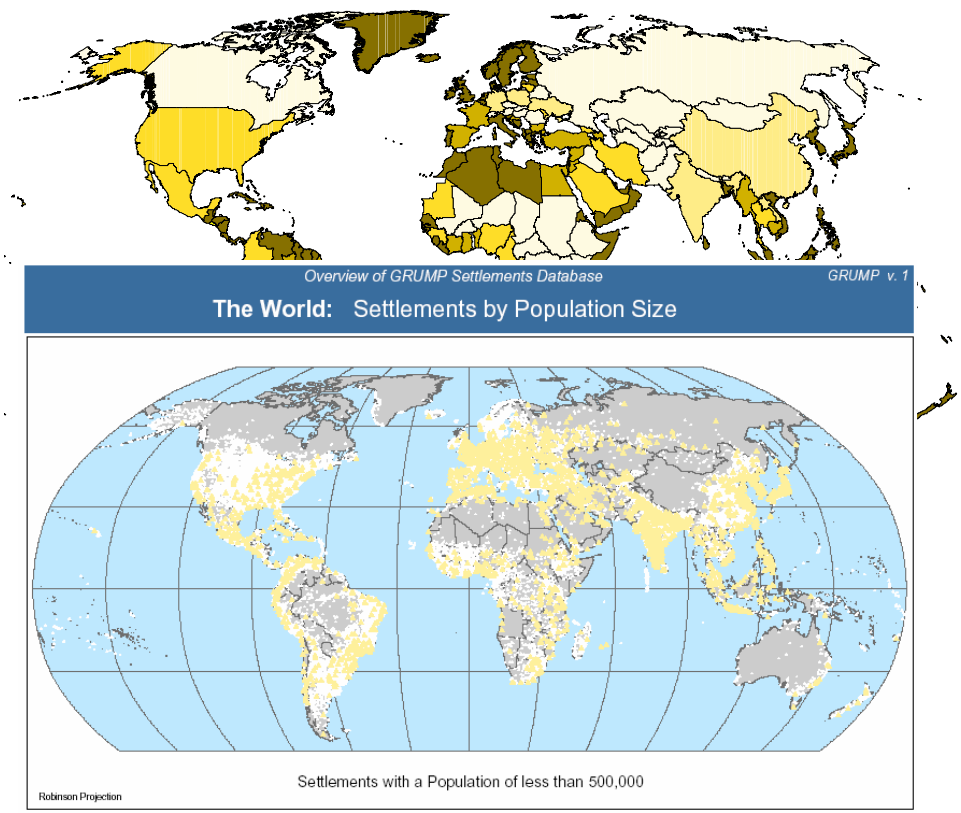
- All shared data, metadata, and products will be made available with **minimum time delay** and **at minimum cost**.

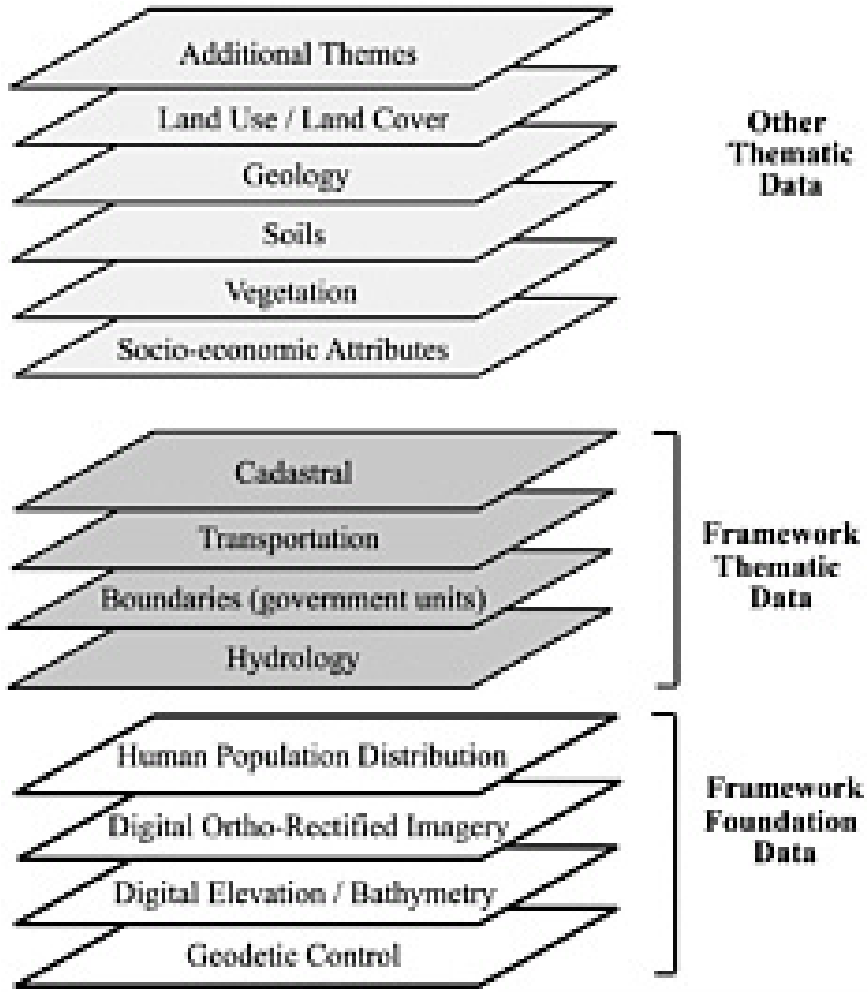
- All shared data, metadata, and products **free of charge or no more than cost of reproduction** will be *encouraged* for **research and education**.

GEOSS 10-Year Implementation Plan, adopted 16 February 2005
(emphasis added)

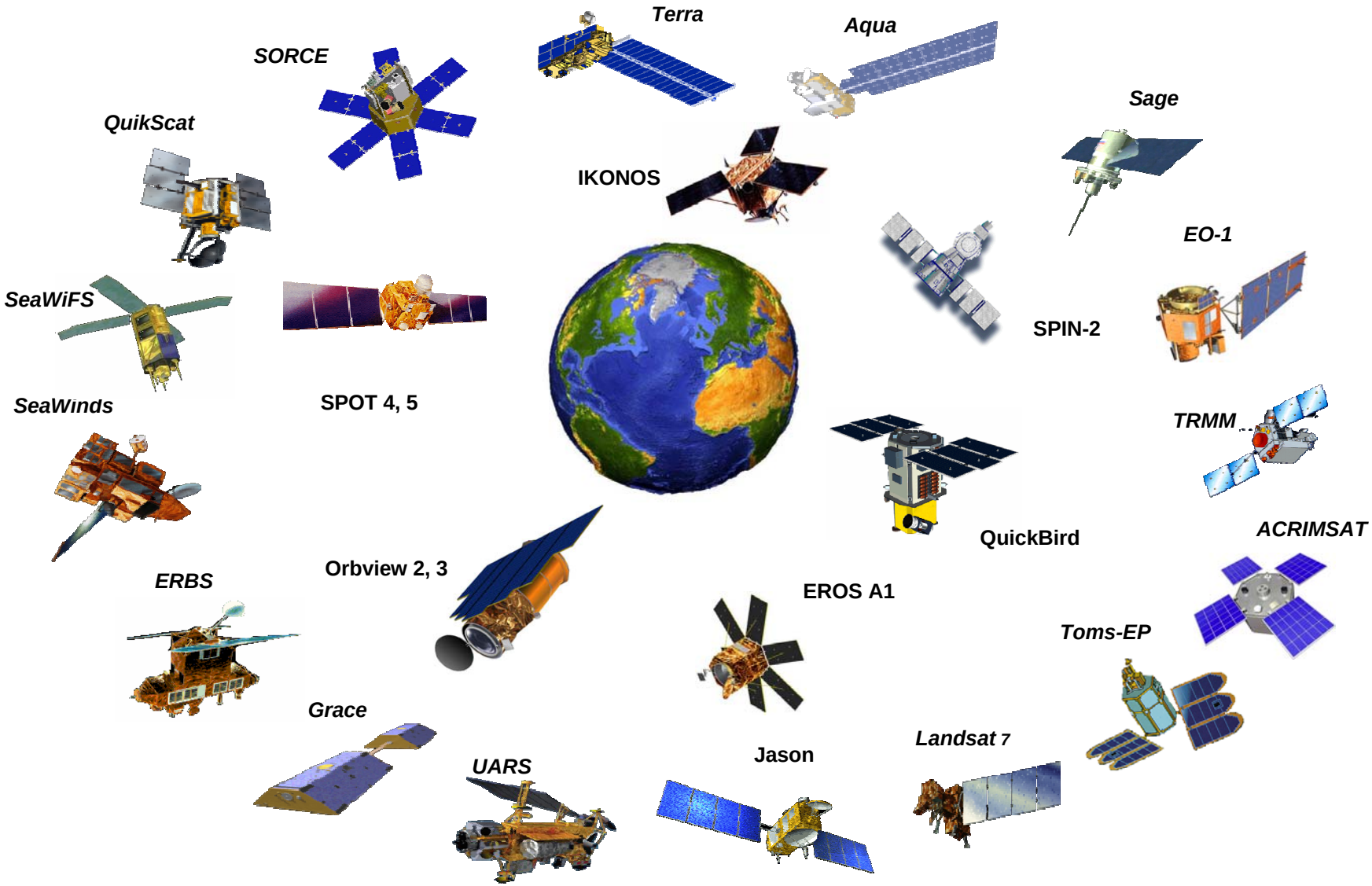
From a "Data Poor" to a "Data Rich" World

- E.g., from data at the national level…
- to detailed subnational data…
- to a rich set of framework and thematic data





Lots of Remote Sensing Data



Support for Decisions and Policy

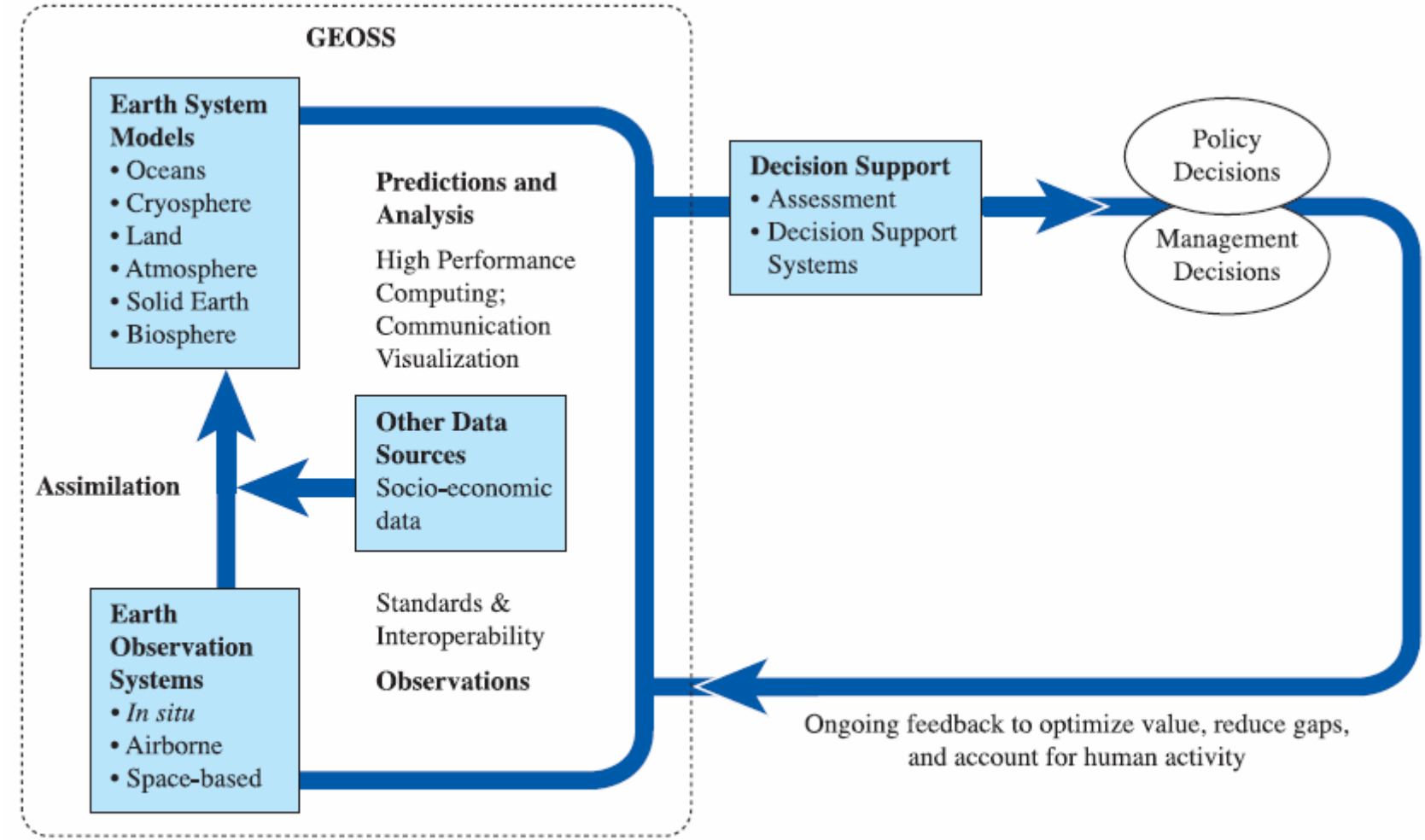
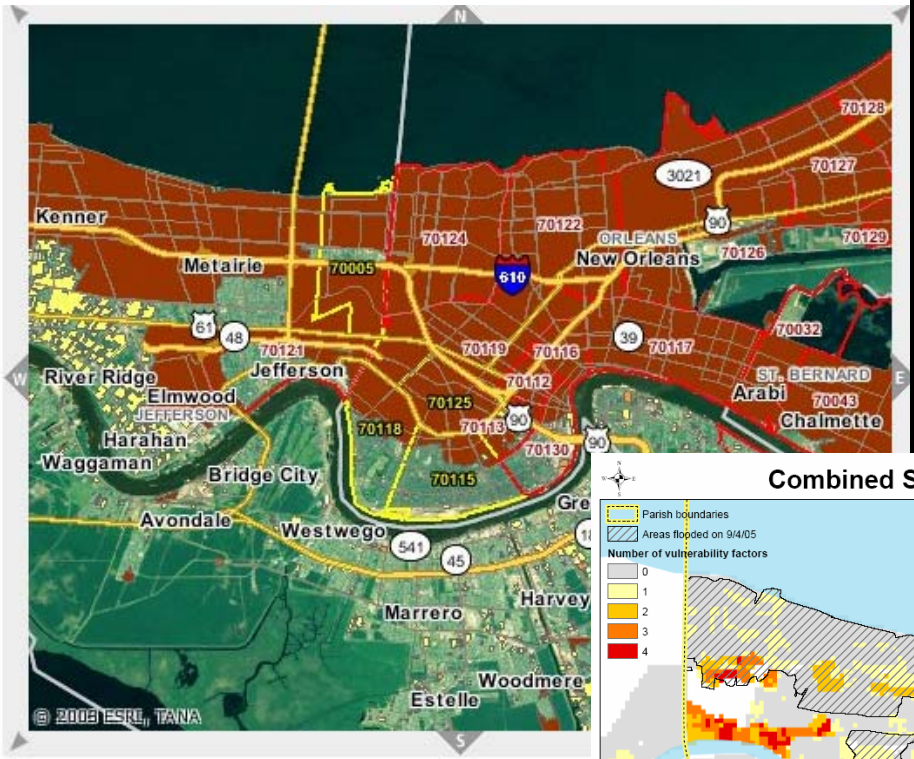
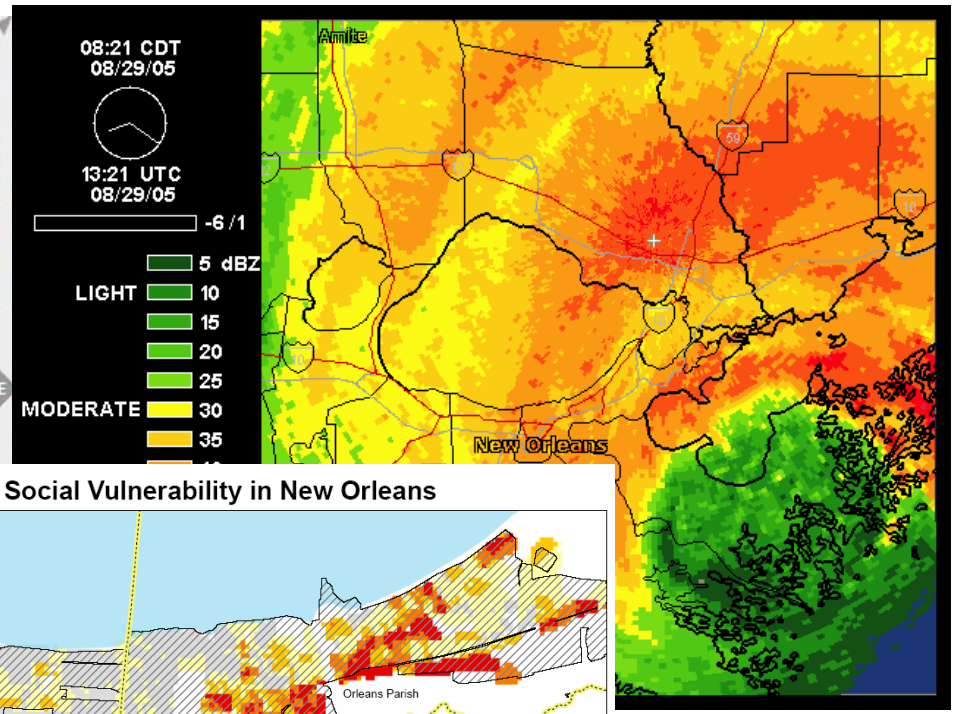


Figure 2.1. This diagram demonstrates the end-to-end nature of data provision, the feedback loop from user requirements, and the role of GEOSS in this process. The primary focus of GEOSS is on the left side of the diagram.

Potential to Save Lives through Real-Time Integration of Data

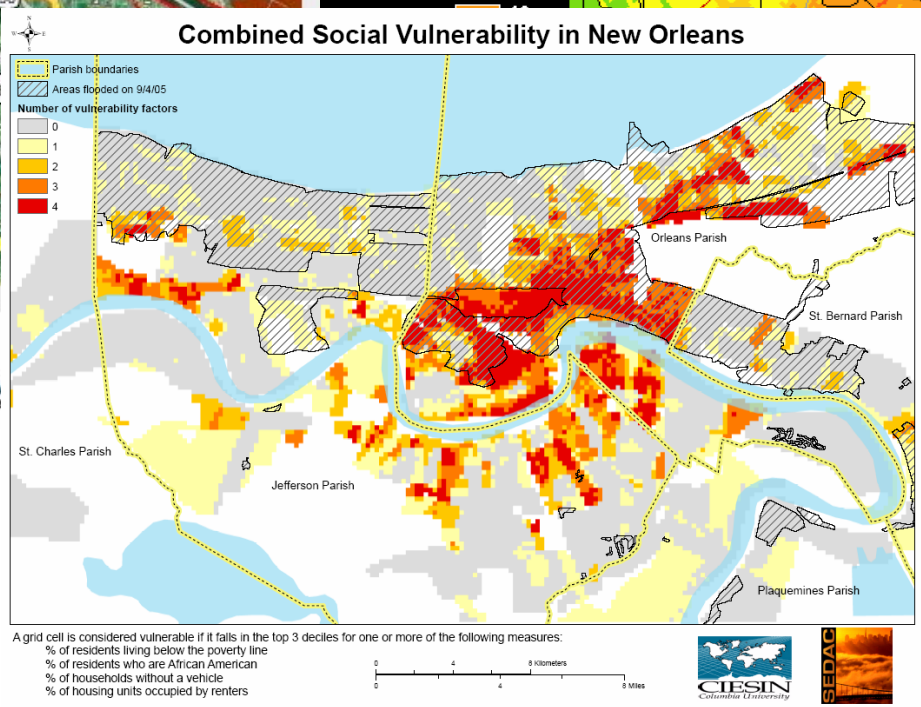




Hazard

Powered by ArcWeb Services

Exposure



Vulnerability

= RISK!

9 GEO Societal Benefit Areas

- Reducing loss of life and property from **natural and human-induced disasters**.
- Understanding environmental factors affecting **human health and well-being**.
- Improving management of **energy resources**.
- Understanding, assessing, predicting, mitigating, and adapting to **climate variability and change**.
- Improving **water resource management** through better understanding of the water cycle.
- Improving **weather information, forecasting and warning**.
- Improving the management and protection of **terrestrial, coastal and marine ecosystems**.
- Supporting **sustainable agriculture** and **combating desertification**.
- Understanding, monitoring and conserving **biodiversity**.

Implementation Issues 1

- Definitions, e.g.:
  - What data and information are covered by the policy?
  - What qualifies as “education and research”?
  - What is the “cost of reproduction”?
  - What exactly is meant by “made available”?
- Exceptions, e.g.:
  - Which international instruments and what national policies and legislation are relevant?
    - National security?
    - Confidentiality and privacy laws?
    - Endangered species protection?
    - Indigenous rights?
    - Intellectual property law?
    - Humanitarian concerns?

IPY Data Policy

…in order to maximize the benefit of data gathered under the auspices of the IPY, the IPY Joint Committee requires that **IPY data, including operational data delivered in real time, are made available fully, freely, openly, and on the shortest feasible timescale**.

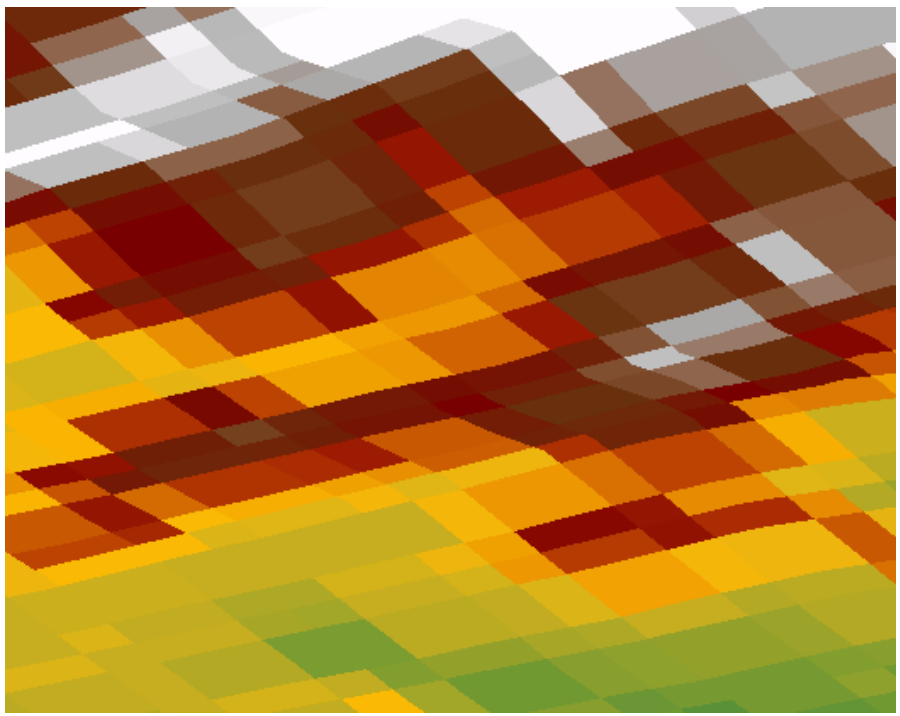
The only exceptions to this policy of full, free, and open access are:

- where human subjects are involved, **confidentiality** must be protected
- where local and traditional knowledge is concerned, **rights of the knowledge holders** shall not be compromised
- where data release may cause **harm**, specific aspects of the data may need to be kept protected (for example, locations of nests of endangered birds or locations of sacred sites).

International Polar Year 2007-2008 Data Policy, 22 May 2006
(emphasis added)

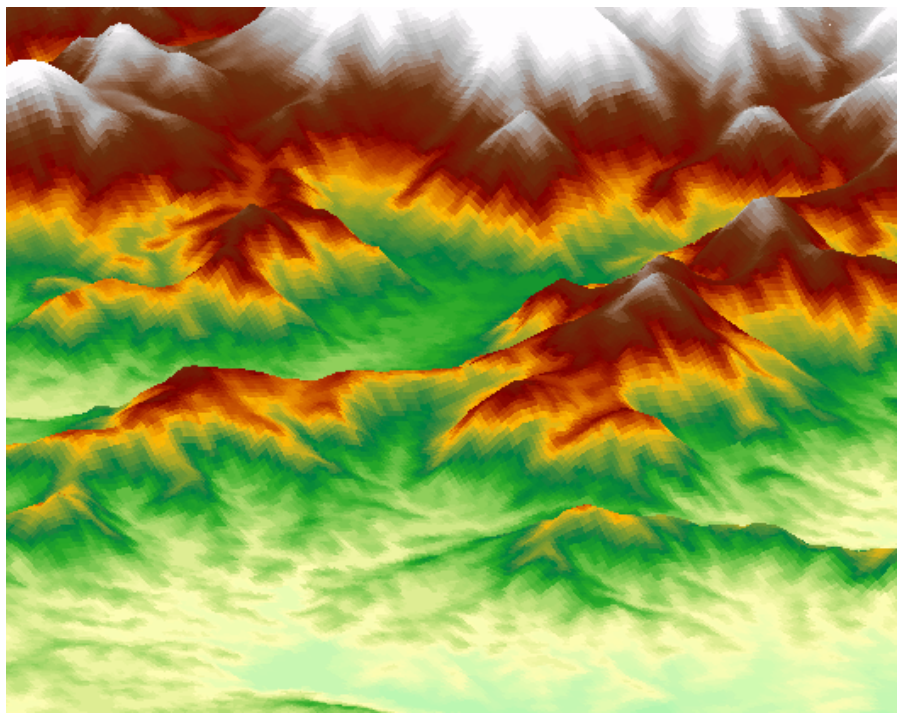
Key Challenge: Data Access

- Increased concerns about national security, intellectual property rights leading to reduced access to data
- Example: U.S. has not released 30-m SRTM for world, despite significant potential benefit for applications
  - Comparison of 30- and 90-m SRTM for Blue Ridge Mountains, VA:



90 m

30 m



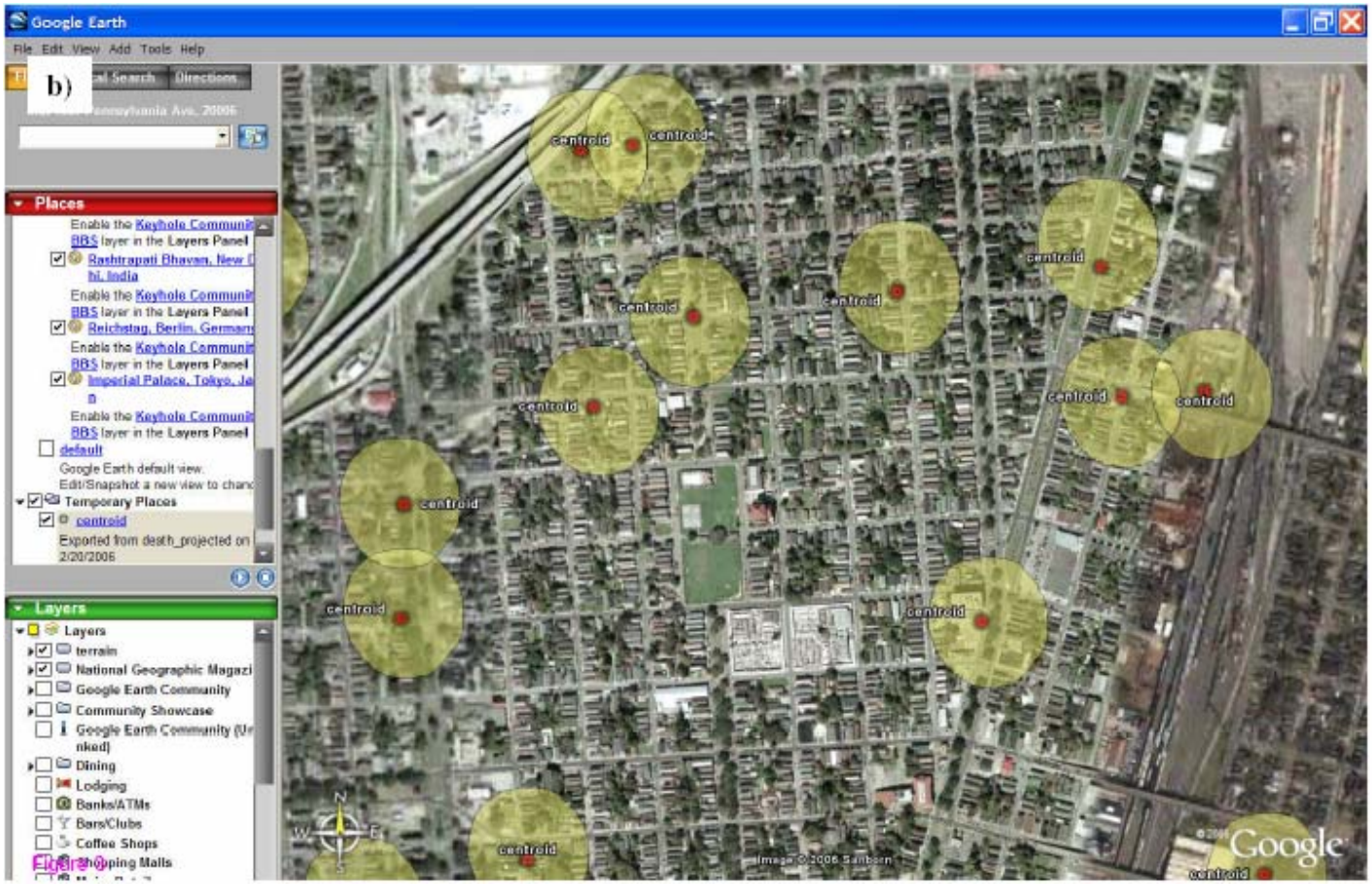
Comparison courtesy of V. Gorokhovich, CIESIN

Key Challenge: Data Withdrawal, Increased IP Restrictions

- Example of Data Withdrawal: U.S. National Geospatial-Intelligence Agency (NGA) has withdrawn its Flight Information Publications (FLIP) and Digital Aeronautical Flight Information File (DAFIF™) from public access due to concerns over copyright claims by other countries.

Key Challenge: Confidentiality

- Example: Use of Google Earth to help pinpoint households where deaths occurred due to Hurricane Katrina, using data from a generalized map



Paper in the *International Journal of Health Geographics* by A.J. Curtis, J.W. Mills, and M. Leitner, "Spatial confidentiality and GIS: re-engineering mortality locations from published maps about Hurricane Katrina" http://www.ij-healthgeographics.com/content/pdf/1476-072X-5-44.pdf

Key Challenge: Protection of Sensitive Resources

- Example: locations of endangered species and the last remaining “wild areas”

Key Challenge: Interoperability

- Example: Inadequate attention to metadata?

CNN.com

sci-tech > space > story page

MAIN PAGE
WORLD
U.S.
LOCAL
POLITICS
WEATHER
BUSINESS
SPORTS
TECHNOLOGY
SPACE
HEALTH
ENTERTAINMENT
BOOKS
TRAVEL
FOOD
ARTS & STYLE
NATURE
IN-DEPTH
ANALYSIS
myCNN

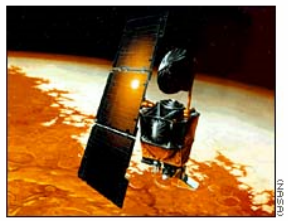
Headline News brief
news quiz
daily almanac

MULTIMEDIA:
video
video archive
audio
multimedia showcase
more services

exploringmars | in-depth specials

NASA's metric confusion caused Mars orbiter loss

September 30, 1999
Web posted at: 1:46 p.m. EDT (1746 GMT)



NASA's Climate Orbiter was lost September 23, 1999

(CNN) -- NASA lost a $125 million Mars orbiter because one engineering team used metric units while another used English units for a key spacecraft operation, according to a review finding released Thursday.

For that reason, information failed to transfer between the Mars Climate Orbiter spacecraft team at Lockheed Martin in Colorado and the mission navigation team in California. Lockheed Martin built the spacecraft.

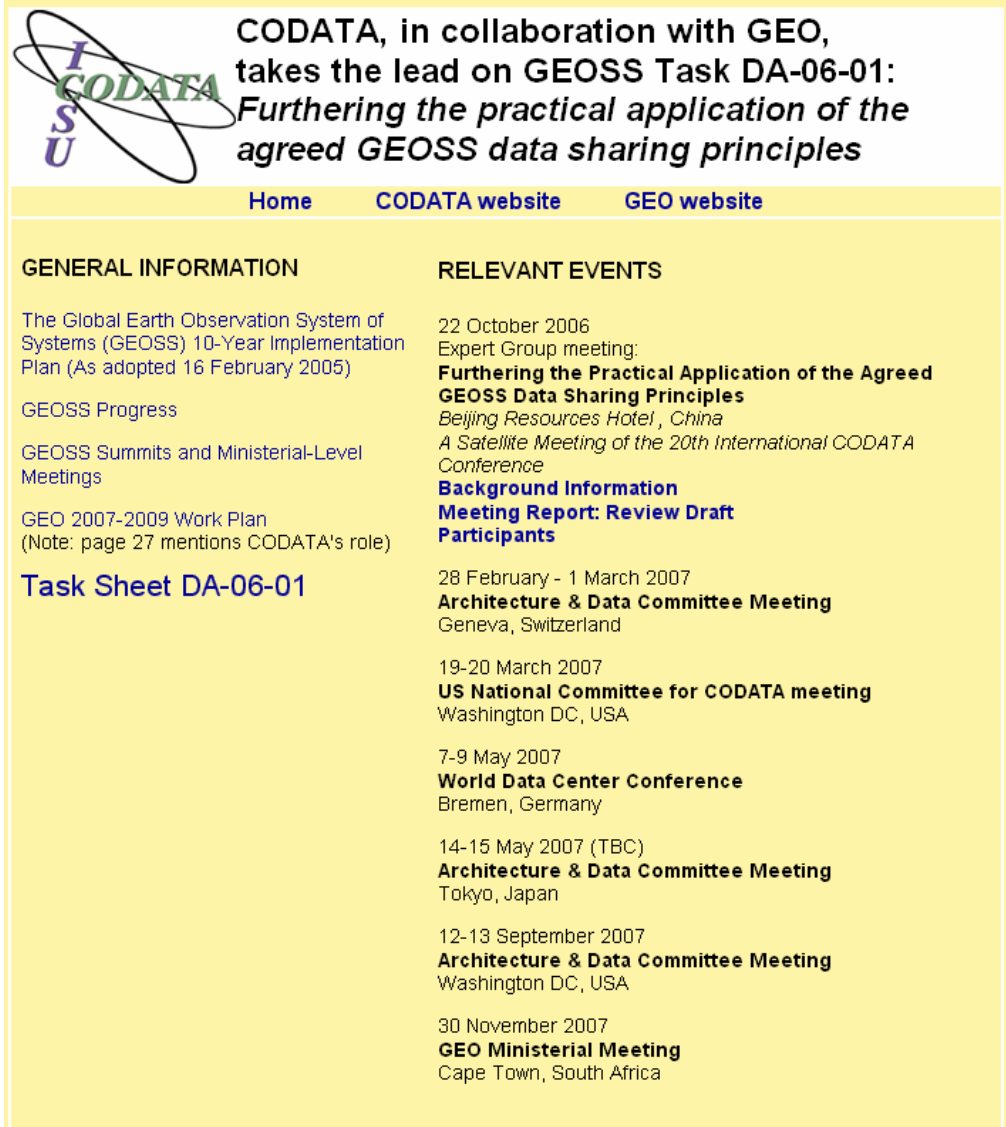
"People sometimes make errors," said Edward Weiler, NASA's Associate Administrator for Space Science in a written statement.

Implementation Issues 2

- Standards, e.g.:
  - How are “minimum time delay” and “minimum cost” determined?
  - What metadata standards need to be met?
- Implementation, e.g.:
  - Use of *ad hoc* or coordinated inter-system agreements?
  - Use of standardized licensing agreements for data and products?
  - Retention of IP through copyright?
  - Use of digital rights management approaches?
- Compliance, e.g.:
  - How would apparent lack of conformance be adjudicated? or negotiated? or publicized?

GEO Task DA-06-01

- Task under Architecture & Data Committee
- Extended into 2007-09 GEO Work Plan
- Initial Experts Meeting held in conjunction with 20th International CODATA Conference in Beijing, China in October 2006
  - More than 30 participants from around the world and from diverse disciplines
  - Meeting report available online
- Task web site established at:
  - http://www.codata.org/GEOSS



CODATA, in collaboration with GEO, takes the lead on GEOSS Task DA-06-01: ***Furthering the practical application of the agreed GEOSS data sharing principles***

Home | CODATA website | GEO website

GENERAL INFORMATION

The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan (As adopted 16 February 2005)

GEOSS Progress

GEOSS Summits and Ministerial-Level Meetings

GEO 2007-2009 Work Plan
(Note: page 27 mentions CODATA's role)

Task Sheet DA-06-01

RELEVANT EVENTS

22 October 2006
Expert Group meeting:
Furthering the Practical Application of the Agreed GEOSS Data Sharing Principles
Beijing Resources Hotel , China
A Satellite Meeting of the 20th International CODATA Conference
Background Information
Meeting Report: Review Draft
Participants

28 February - 1 March 2007
Architecture & Data Committee Meeting
Geneva, Switzerland

19-20 March 2007
US National Committee for CODATA meeting
Washington DC, USA

7-9 May 2007
World Data Center Conference
Bremen, Germany

14-15 May 2007 (TBC)
Architecture & Data Committee Meeting
Tokyo, Japan

12-13 September 2007
Architecture & Data Committee Meeting
Washington DC, USA

30 November 2007
GEO Ministerial Meeting
Cape Town, South Africa

Expected Task Outputs

1) “White” Paper on Guidelines for Implementing the GEOSS Data Sharing Principles
   - Writing team to include Joanne Gabrinowycz, University of Mississippi and Dave Clark, NGDC
   - Workshop later this year to develop and agree on text
2) New language on data policy in the Declaration planned for the November 2007 Ministerial Summit in Cape Town, S. Africa
   - “Finally, the Declaration should address data sharing principles. Developing more open data policies and data sharing agreements in Earth Observation remains a priority for several GEO Members. **The Declaration could identify specific goals to reduce data policy barriers to align with the GEO goal of ‘free and open exchange’ by a target date**.”

Draft White Paper Outline 1

I. Executive Summary

II. Background

III. Review of past experience with data sharing principles & policies

a) Other international scientific/Earth Observation programs
b) Relevant national/regional data sharing principles & policies
c) Review of studies of the value of open data access

IV. Options for implementing the GEOSS data sharing principles

a) Alternative approaches
b) Key issues surrounding exceptions to the principles
c) Incentives for compliance with the principles
d) Monitoring, enforcement, and appeal processes

Draft White Paper Outline 2

V. Case studies ("high payoff" activities tested against the alternative approaches)

a) Access to real-time and historic GEOSS data for rapid humanitarian response
b) Research use of integrated GEOSS data for climate change impacts assessment
c) Educational use of multidisciplinary GEOSS data and information products in developing countries
d) Public/private sector use of real-time and near-real time GEOSS data for weather forecasting
e) Private sector use of new GEOSS sensors for new benefit areas, e.g., renewable energy
f) Local government use of high-resolution GEOSS data for biodiversity conservation and/or coastal/marine ecosystem management
g) National government use of time series GEOSS data for regional water resource management and/or combating desertification
h) Public health use of specialized GEOSS data for infectious disease management
i) General public use of GEOSS data for 3- and 4-dimensional mapping and visualization

VI. Technical Implementation Issues

a) Data interoperability
b) Data quality
c) Data standards
d) Low-cost access
e) Digital rights management
f) Metrics

Draft White Paper Outline 3

VII. Policy Implementation Issues

a) Engaging stakeholders
b) Promoting the open access “ethos”
c) Encouraging harmonization
d) Supporting transparency
e) Ensuring sustainability
f) Establishing meaningful goals and target dates

VIII. Recommended Guidelines for GEOSS Data Policies

a) Recommended definitions of terms
b) Recommendations regarding what international instruments are relevant, how they should be applied to the GEOSS principles, and whether consideration should be given to changing any of them (if possible).
c) Recommendations regarding what existing national policies and legislation are relevant (in general), how they should be applied to the GEOSS principles (in general), and whether consideration should be given to changing any of them (if possible).
d) Recommendations regarding specific GEOSS data policies
e) Recommendations regarding coordination with other GEO tasks
f) Recommendations regarding specific goals and target dates

IX. Summary and Conclusion

November 2007 Ministerial Summit

- 2 Working groups established
  - WG1 on declaration text: R. Chen nominated by OGC and ICSU
  - WG2 on GEOSS accomplishments: G. Percival of OGC and G. Glaser of ICSU
  - First meetings in April
- Possibility of side event(s) in conjunction with Summit to address data policy issues

Contributions Welcome!

- Inputs on data policy issues
- Participation in data policy workshop (August-September time frame)
- Authors for specific sections of the White Paper
- Specific suggestions for Declaration text
- Coordination with other GEO tasks
- Coordination with ADC and other GEO Committees
- Help in planning side event(s)
- Travel and other financial support

Thanks to:

- Nancy Colleton, The Alliance for Earth Observations
- Paul Uhlir, US National Committee for CODATA

Thanks for attending!

Background information available at:

- http://www.codata.org/GEOSS