The development of international data policy for science

Ray Harris
University College London

CODATA'06 Conference
Beijing, China
25 October 2006

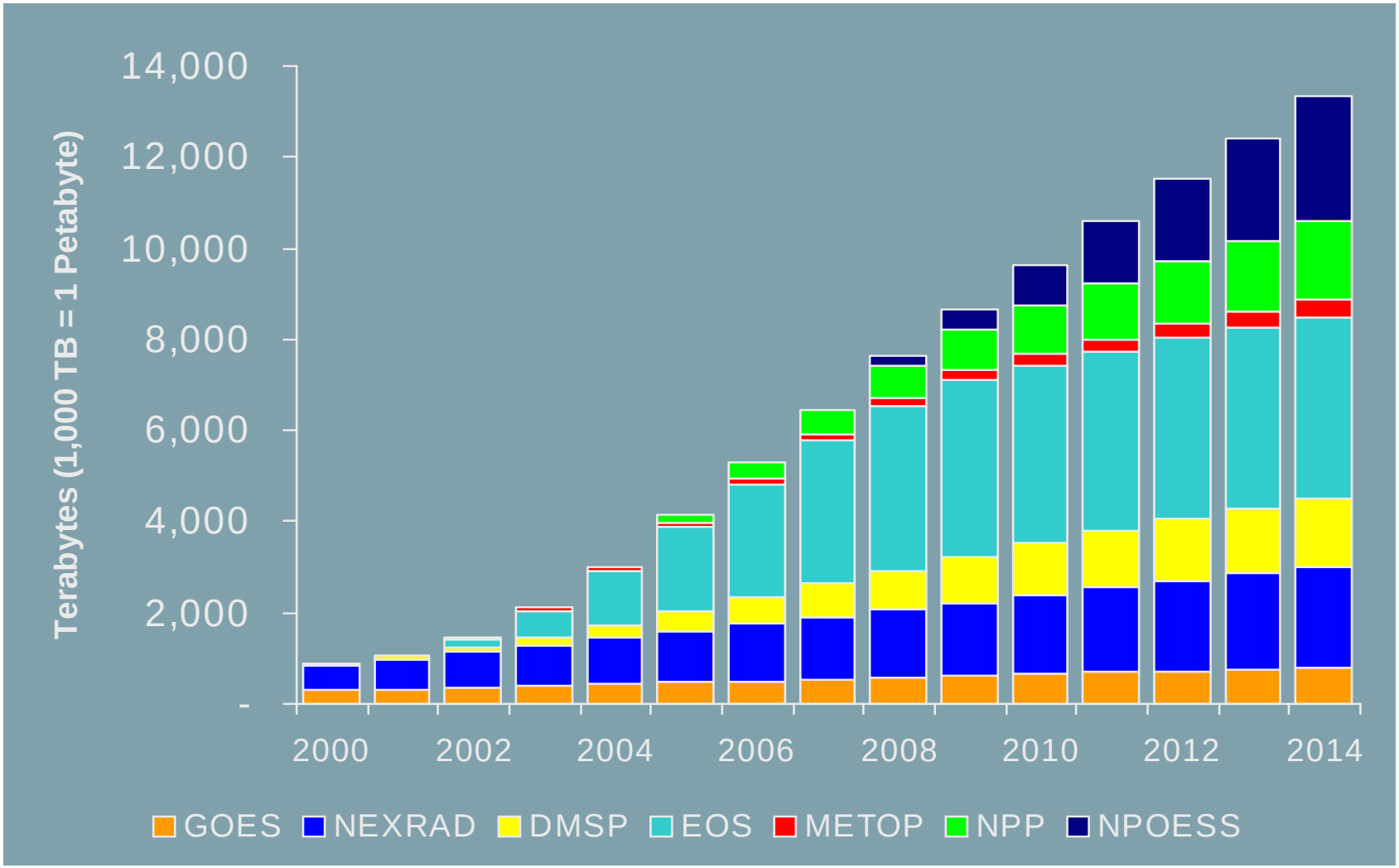
Agenda

- Change in science
- Technology strides
- World concern for information
- ICSU panel on data and information
- International Polar Year data policy

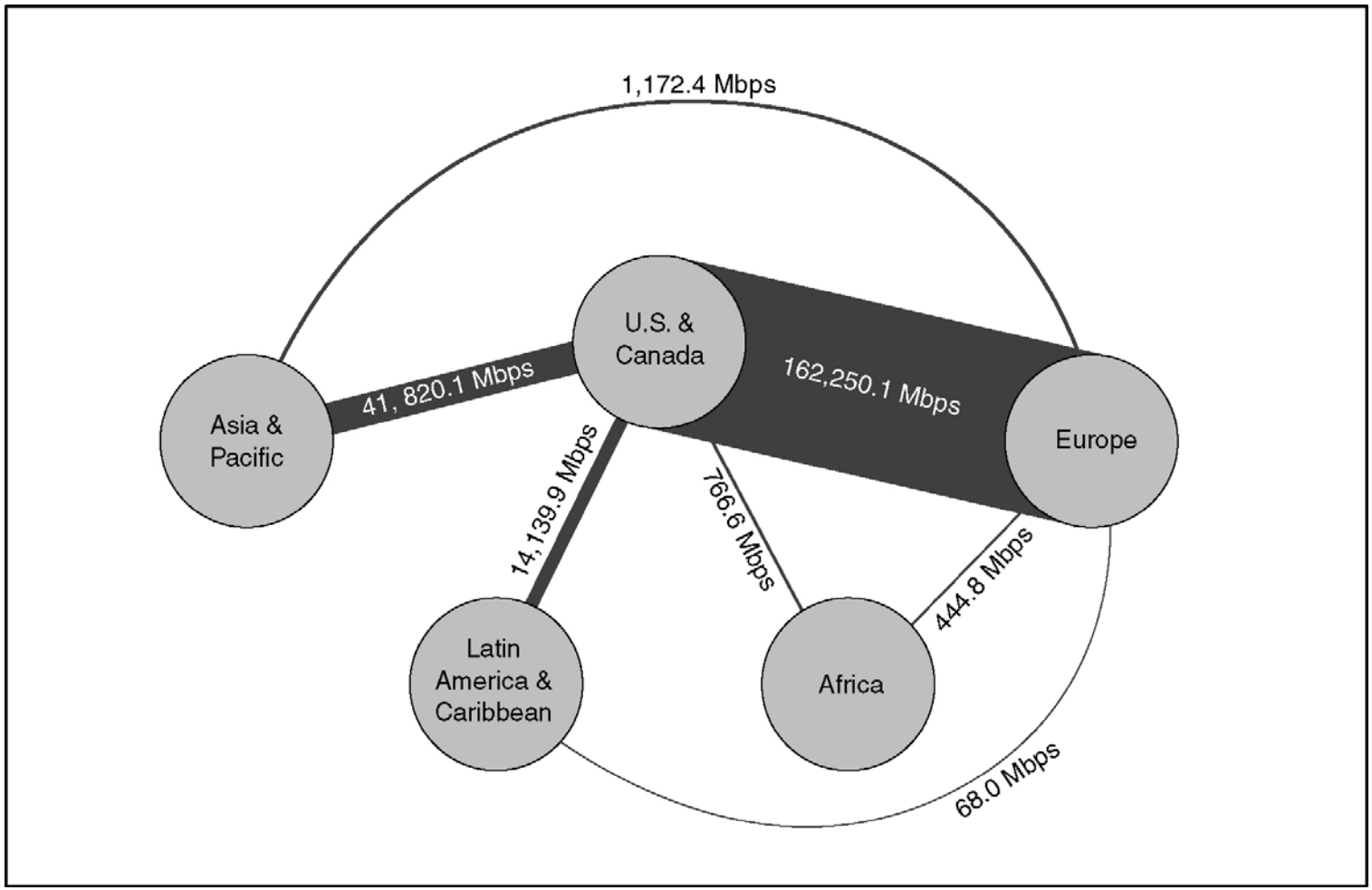
Context of change in science

- Data acquisition capabilities
- Computational capacity and connectivity
- Distribution and management; data sharing
- Digitisation : data, publications, pictures, voices, art, traditional and indigenous knowledge
- New possibilities, e.g. human genome, monitoring the whole of planet Earth

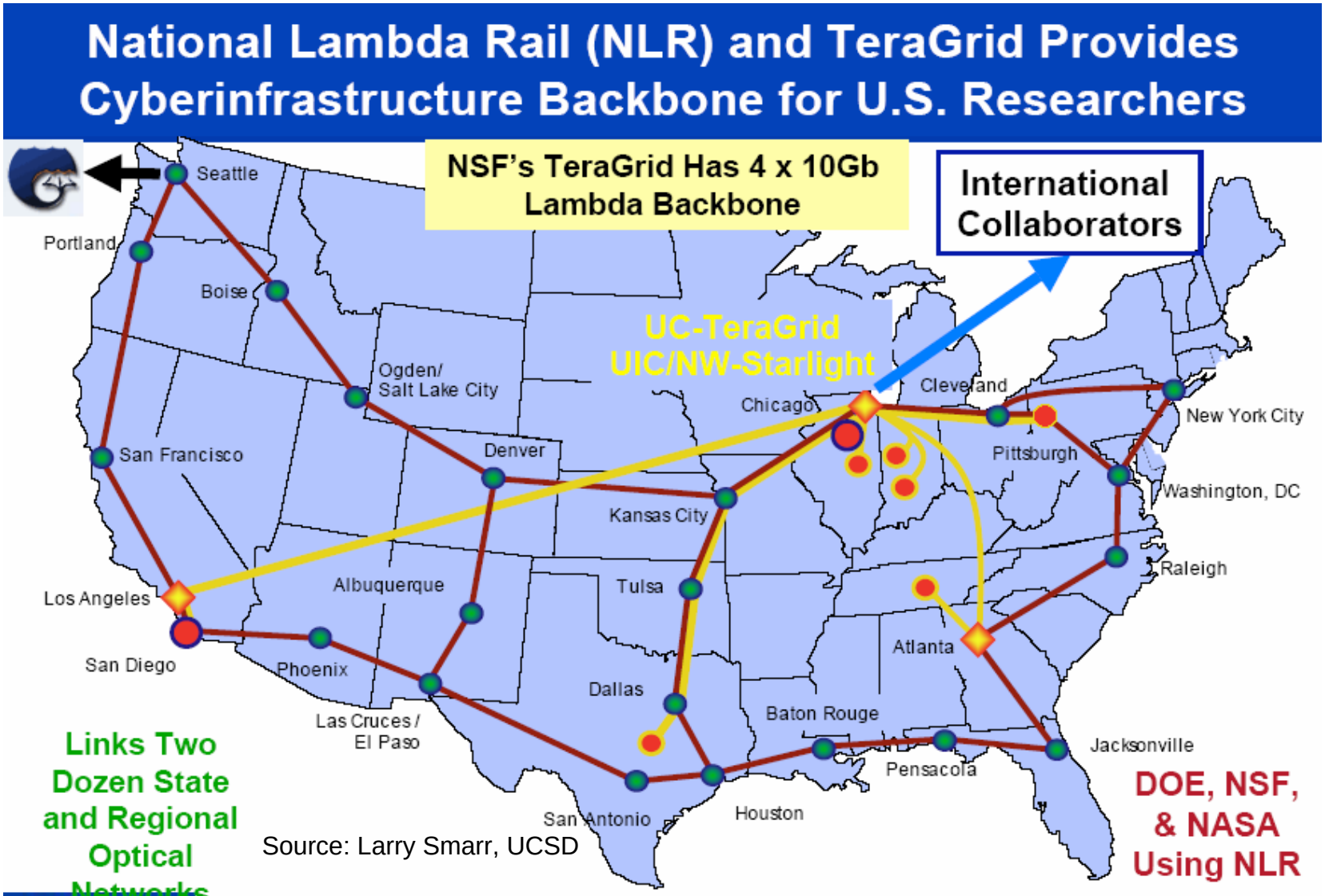
US NOAA data archive growth



Digital divide – data rates



US National Lambda Rail (NLR) and TeraGrid



Source: Larry Smarr, UCSD

Global connections

Links university research centres
10 Gbps

Source: Larry Smarr, UCSD

World Summit on the Information Society WSIS

- Geneva 2003 Phase I 2003
- Tunis November 2005 Phase II (*N.B. Africa*)
  - 46 Heads of State and Government, Crown Princes and Vice-Presidents and 197 Ministers/Vice Ministers and Deputy Ministers
  - 5,857 participants representing 174 States and the European Community
  - 1,508 participants representing 92 international organizations
  - 6,241 participants representing 606 NGOs and civil society entities
  - 4,816 participants representing 226 business sector entities
  - 1,222 accredited journalists from 642 media organizations of which 979 onsite from TV, radio, print and online media worldwide

Tunis Agenda for the Information Society

- “the Internet, a central element of the infrastructure of the Information Society, has evolved from a research and academic facility into a global facility available to the public”
- Task Force on Financial Mechanisms
    - Voluntary Digital Solidarity Fund
- Internet governance : shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet
- Follow up action line : Ethical dimensions of the Information Society

International Council for Science ICSU

- Panel Area Assessment on Data and Information
- Panel Area Assessment on Environment in Relation to Sustainable Development
- Panel Area Assessment on Capacity Building
- Period 2003 – 2005
- Prepare for the ICSU General Assembly, Suzhou, China, October 2005 and the ICSU strategic plan 2006 - 2012

ICSU Panel Area Assessment on Data and Information

- Terms of reference
  - Mission and role for ICSU
  - Strategic framework for ICSU for 5 – 10 years
  - Propose changes in collaboration and coordination within ICSU and with outside organisations
  - Examine ethical and wider policy issues

International Council for Science ICSU Panel Area Assessment on Data and Information

- Roberta Balstad, USA, chair
- Jean Bonnin, France
- Marc Brodsky, USA
- Liu Chuang, China
- Carlos Correa, Argentina
- Norihisa Doi, Japan
- Ray Harris, UK
- Andrew M. Kaniki, South Africa
- Vitaly Nechitailenko, Russia
- Pierre Ritchie, Canada, ICSU
- T.B. Rajashekar, India

- Science
- Social science
- Information management
- Publishing

Issues considered

- Scope of data and information
- Role of the public sctor
- Role of the private sector
- Professional data management
  - Archiving
  - Data rescue
  - Metadata
  - Standards
  - Interoperability
- Scientific publications and open access publishing
- Systems for data dissemination
- Equitable access
- Who pays?
- Intellectual property rights
- Digital divide
- Organisation within ICSU

Main recommendations

- Science is best served by minimal restraints
- ICSU should assume an international leadership role in identifying and addressing critical policy and management issues related to scientific data and information
- Promote professional data management
    - consistency
    - quality – metadata and standards
    - permanent preservation
    - funding should be routine
- Balance of considerations on publications
    - open access publications but may lead to author pays
    - funding mechanism for learned societies
    - equitable access to on line journals, especially in developing countries

Main recommendations

- Who pays?
- Data production and management are costly
- Ensuring the long-term accessibility of increasing quantities of scientific data and information will necessitate increased public (and private) investment in data management and long-term institutional support
- ICSU and its members should explore various solutions to meet the financial challenge of providing full and open access to scientific data and universal and equitable access to publications

Main recommendations

- Ethical concerns
  - Individual and population data
  - Cross-referencing of data sets
- Intellectual property rights
  - Becoming more important, e.g. WIPO, WTO
  - Danger of restrictive use conditions
  - Public and private data sets are beneficial to science

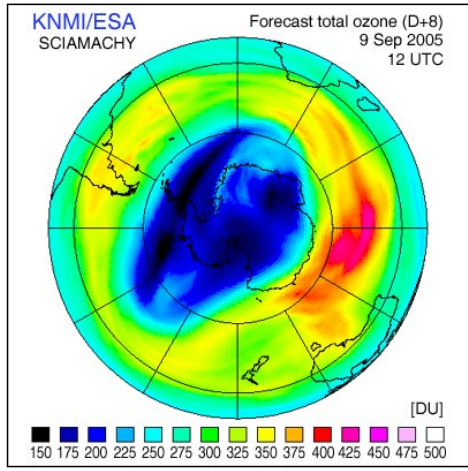
Recommendations to CODATA

- Long term strategy on international data management
- Improve lines of communication with ICSU and its other bodies
- Encourage more ICSU members to participate in CODATA
- Links with the World Summit on the Information Society

ICSU strategic plan 2006 - 2012

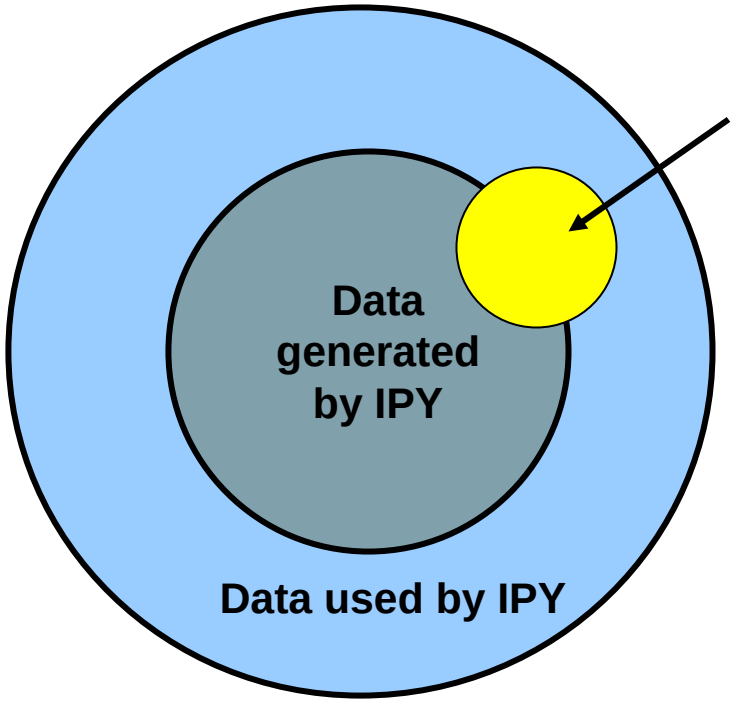
- Universality of science:
  - freedom of movement, association, expression and communication for scientists
  - equitable access to data, information and research materials
  - against discrimination on ethnic origin, religion, citizenship, language, politics, gender, sex, age

International Polar Year (IPY) data policy

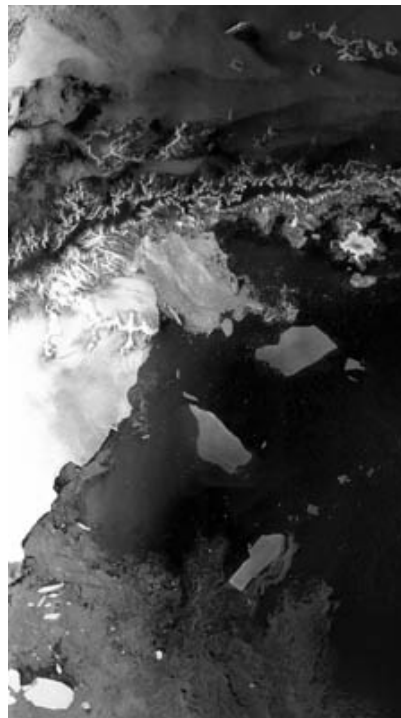


- Consistent handling of data
- Balance
  - Widespread access
  - Rights of investigators
  - Rights of indigenous peoples
- IPY data, including operational data delivered in real time, should be made available fully, freely, openly and on the shortest possible time scale
- Exceptions
  - Human subjects
  - Local knowledge holders
  - Where data may cause harm, e.g. bird nesting sites

Concept



Special cases
•Human subjects
•Intellectual property
•Possible harm



Conclusions

- Data policy is a servant of objectives
- A challenge is to make objectives more explicit
- Science needs minimal restraints
- Data and information collection is costly
- Professional data management is essential