Future Earth: Data management challenges

Sandy P. Harrison
University of Reading, UK
Macquarie University, Australia
Future Earth is the global research platform providing the knowledge and support to accelerate our transformations to a sustainable world.
The origins of Future Earth
The remit of Future Earth

To provide the knowledge required by societies to face the risks of global environmental change and to seize opportunities in a transition to global sustainability

• Three key elements: scientific integration, co-production of knowledge, truly global participation

• Three research themes: dynamic planet, global development, transformation towards sustainability

Source: Mauser et al., COSUST, 2013
The Governance of Future Earth

The Science and Technology Alliance for Global Sustainability

Governance Council

Engagement Committee

Science Committee

Executive Secretariat

Transformations towards sustainability
Disciplinary science Projects Integration & Synthesis

Dynamic Planet

Global Development
Dynamic Planet

- Approaches and Models
  - drivers
  - societal system
  - observing
- States and Trends
  - explaining
  - thresholds
  - understanding
- Critical Zones
  - coasts
  - tropical forests
  - polar regions

Global Development

- Stewardship of resources
  - clean air
  - mining
  - biodiversity
- Ecosystem services
  - Trade-offs
  - climate change
  - fisheries
- Equitable access
  - food security
  - water availability
  - healthy environment

Transformations towards Sustainability

- Transformation process
  - economy
  - mega-cities
  - development options
- Innovation and ideas
  - trade-offs
  - emerging technology
  - assessment of policies
- Global and regional governance
  - international law
  - incentives
  - regional enforcement
All research is data driven

Data are the major legacy of any science programme

Appropriate data management has to be core to the design of any programme
Data challenges facing Future Earth

Integrated and interdisciplinary datasets, indicators, visualizations, scenarios, information...

Existing Data
Existing Knowledge
New Observations
Models

Social, economics, ecosystems, geophysics
Data challenges facing Future Earth

Future Earth will need access to data and will bring large volumes of diverse environmental, biological and social data together. As observing, surveying and modelling systems become more complex, the challenge of accessing and bringing large volumes of diverse data together increases. Future Earth will depend on international initiatives, such as the ICSU World Data System...ICSU-CODATA can make an important contribution on the policy side of scientific data management, especially promoting open environments for data.
Summary of data recommendations for Future Earth

1. Policies for the management, preservation, and dissemination of scientific data collected under the auspices of Future Earth should be developed in collaboration with CODATA and WDS. These policies must be in place at the onset of the programme. All data policies should be accessible to all participants on the Future Earth website.

2. Future Earth data policies should include a mandatory requirement for archiving data from Future Earth research projects in a funded data archive and for listing the metadata in a metadata catalogue linked to the Future Earth web page.

3. Insofar as possible, existing data centres and archives should be used for Future Earth data. New data centres should be established in connection with the ICSU Regional Centres. Training should be made available for those who work in the centres.

4. If a scientist has no archive immediately available for research data, the ICSU World Data System should be responsible for finding an “archive of opportunity” for the scientist to use in depositing data.

5. Funding for data documentation and transfer to data centres should be in place at the beginning of the Future Earth research programme and should be supported as a routine part of research project funding.
6. Data should be documented and archived within one year of the completion of the research project.

7. All data collected or modified for research under Future Earth auspices should be declared to be co-owned by the investigator and by Future Earth. These data should be made available to all who request them on an equitable basis and at no more than the cost of reproduction.

8. Provision should be made for a web site providing separate data access by public and private sector users. This web site should provide non-scientific users with guidance to the databases and links to Future Earth research.

9. Data integration should be a recognised research task eligible for funding in the Future Earth research programme.

10. Data needs for each major area of Future Earth research should be discussed early in the planning process, in collaboration with CODATA and WDS. Where deemed useful, needed baseline data sets should be identified.

11. It is also essential to recognise that effective global science will require the development of data access and management capacity in developing countries, the training of potential users, especially non-scientists, about the data, and the identification of adequate financial support for data activities. CODATA and WDS can assist in identification of institutional partners and collaborators in this arena.
Pious statements and reality

• Future Earth (WDS, CODATA) cannot force compliance
• Data issues are not high on the Future Earth agenda; focus is on “science”
• There is no “data champion” on the Future Earth SC or EC
• There is as yet no consideration of data quality or the interface with data users
• Meshing data sets is a lot of work
Huge volumes of data

- CMIP5 generated 1380 TB of data
- An order of magnitude more than CMIP3
- Only 16% of the model outputs generated
- CMIP6 already contemplating a 5-fold increase in number of experiments, with more outputs to be archived
- This is a single “data source”
Moving targets and data quality
Moving targets and data quality

Who will be the guide?
Same old, same old

Federation of projects that currently exist under the old banners (IGBP, Diversitas, IHDP, WCRP) .... Have to be formally approved as FE

FROM A DATA MANAGEMENT PERSPECTIVE: danger of bringing the old baggage with them
What are WDS and CODATA doing?

- WDS and CODATA have a mission to support ICSU Programmes with their data management activities.
- Created draft Future Earth Data Sharing Principles.
- Advocating a data expert within both the Future Earth Scientific Committee and Secretariat.
- Consulting with the affiliated projects.
- Promoting explicit mechanisms for governance and implementation of data management and dissemination.
- Working with funders to ensure recognition of need for data archiving support.
And the bottom line …

- Data managers need to get involved in Future Earth right from the start (directly or through affiliated projects)
- Data generators need to take archiving and documentation seriously
- Funders need to take archiving and documentation seriously
- Future Earth needs to a Data Champion or Data Team to address issues of archiving, qc, access, meshing, documentation
- WDS and CODATA must assist with all of this
FAILURE IS NOT AN OPTION

THANKS!