



Future Earth Biodiversity Symposium  
Yale  
May 3-6, 2015

# CODATA, Future Earth, and Biodiversity

**John Broome**  
**Treasurer, CODATA**

**futureearth**  
research for global sustainability



**ICSU**  
International Council for Science

# CODATA Strategy: Exploiting the Data Revolution

New CODATA President  
Geoffrey Boulton, FRS  
Chair of *Science as an  
Open  
Enterprise* Report



New CODATA Executive  
Committee elected at  
GA in New Delhi, Nov  
2014



Three priorities:

- 1. data policies:** supporting implementation of data principles and practice
- 2. data science:** addressing the frontier issues of data science
- 3. data education:** capacity building



# CODATA Activities ....

## International Data Week Sept. 2016 (DC-USA)

- International conference on data science.
- Brings together CODATA, WDS & RDA.

## Relaunched *Data Science Journal*.

- New Editor-in-Chief, Dr. Sarah Callaghan
- Partnership with a dynamic OA publisher.

## Data Education

- Framework for Data Science Summer Schools
- Data Science Capacity Initiative

## Data Frontiers

- Uniform Description System for Nanomaterials
- Big Data and Data Integration for International Science.
- Task groups





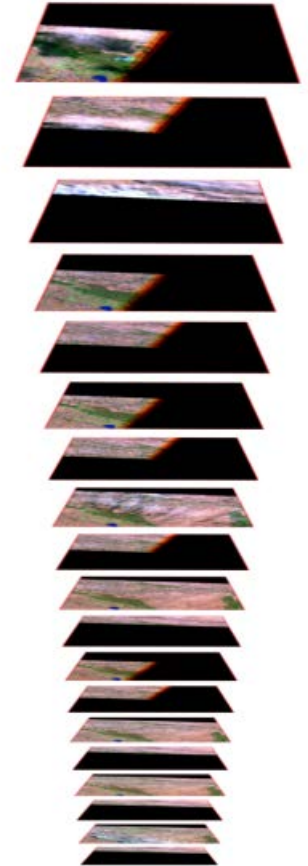
# Data Principles for Future Earth

Proposed by CODATA and WDS in 2013:

- **Excellence in Data Management:** data generated and modified in Future Earth, and associated research products—such as code—will be managed throughout the research lifecycle in accordance with good practice.
- **Openness and Protection:** data, and other research products, generated and modified in Future Earth will be made as openly available as possible, with minimum delay and at minimum cost, while respecting relevant international agreements, national policies and legislation for the protection of personal, sensitive and commercial data;
- **Integrity and Legacy:** data, and other research products, generated and modified in Future earth will be discoverable, accessible, intelligible and reusable, in the short and long term, and will therefore be selected appropriately, quality-assessed, furnished with appropriate metadata, machine readable licenses, and maintained in trusted digital repositories.

# Requirements for Data Intensive Science

- Interdisciplinary science requires integration and analysis of increasingly large and complex data - often by non-specialists – making data standardization and documentation essential.
- Data portals, downloads and local storage and processing will become inadequate and inefficient.
- The future will require “platform approaches” combining HP computing, data storage, standard data services, and analysis based on standard models, variables.



# Moving Forward ....

- Experiences with geological data : OneGeology



- Based on openness and accessibility for common good.
- Approaches: Bottom-up in Canada failed - Top-down global was successful.
- Implemented using federated model with a standard data structure and semantics.

- Questions for Biodiversity Monitoring Community

- Is the “data culture” open and sharing ?
- Are access and use policies standardized and accepted?
- Are standards documented for data, metadata, services, models, variables ?
- Would the community establish and sustain a single biodiversity platform providing access to both EO and conventional monitoring data ?



