# CODATA and Data Diplomacy

Simon Hodson, Executive Director, CODATA

www.codata.org

International
Science Council

# CODATA Prospectus:

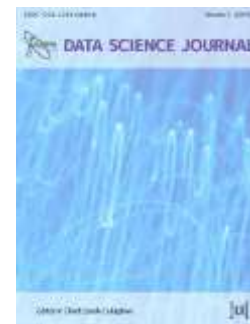## https://doi.org/10.5281/zenodo.2559274

## Principles, Policies and Practice



## Frontiers of Data Science



### Data Science Journal

CODATA 2017, Saint Petersburg 8-13 Oct 2017

## Capacity Building

# Major CODATA Initiatives

- **Data Policy Committee: http://www.codata.org/strategic-initiatives/international-data-policy-committee**

- **Data Interoperability and Integration for Interdisciplinary Research: http://dataintegration.codata.org**

- **African Open Science Platform: http://www.codata.org/strategic-initiatives/african-open-science and https://doi.org/10.5281/zenodo.2222418**

- **CODATA-RDA Schools of Research Data Science: http://www.codata.org/working-groups/research-data-science-summer-schools**



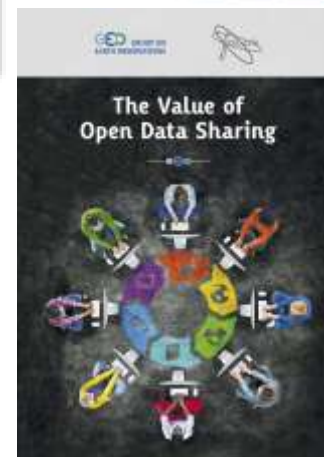THE **AFRICAN OPEN SCIENCE PLATFORM**

**The Future of Science and Science for the Future**

# CODATA Data Policy Reports Since 2015

- Current Best Practice for Research Data Management Policies (for Danish e-Infrastructure Cooperation, 2015)
http://dx.doi.org/10.5281/zenodo.27872

- The Science International Accord on Open Data in a Big Data World (for ICSU, 2015)
http://www.science-international.org/#accord

- The Value of Open Data Sharing (for GEO, 2015)
http://dx.doi.org/10.5281/zenodo.33830

- Guidelines for the Legal Interoperability of Research Data (with RDA, 2016)
https://doi.org/10.5281/zenodo.162241

- Business models for sustainable research data repositories (with OECD, 2017)
https://doi.org/10.1787/302b12bb-en

- European Commission Expert Group, Chaired by Simon Hodson, Turning FAIR into Reality (2018)
https://doi.org/10.2777/1524

# Why Open Science / FAIR Data?

- **Good scientific practice depends on communicating the evidence.**

    - Open research data are essential for reproducibility, self-correction.

    - Academic publishing has not kept up with age of digital data.

    - Danger of an replication / evidence / credibility gap.

    - Boulton: to fail to communicate the data that supports scientific assertions is malpractice

- **Open data practices have transformed certain areas of research.**

    - Genomics and related biomedical sciences; crystallography; astronomy; areas of earth systems science; various disciplines using remote sensing data…

    - **FAIR data helps use of data at scale, by machines, harnessing technological potential.**

    - Research data often have considerable potential for reuse, reinterpretation, use in different studies.

- **Open data foster innovation and accelerate scientific discovery through reuse of data within and outside the academic system.**

    - Research data produced by publicly funded research are a public asset.

# Policy Push for Open Research Data

- Bits of Power: issues in global access to scientific data, NAS, 1997
- The three Bs (Budapest, Berlin and Bethesda) and Open Access, 2002-3
- **OECD Principles and Guidelines on Access to Research Data, 2004, 2007**
- UK Funder Data Policies, from 2001, but accelerates from 2009
- NSF Data Management Plan Requirements, 2010
- Royal Society Report 'Science as an Open Enterprise', 2012
- India National Data Sharing and Accessibility Policy, 2012 and 2014.
- OSTP Memo 'Increasing Access to the Results of Federally Funded Scientific Research', Feb 2013
- **G8 Science Ministers Statement, June 2013**
- **G8 Open Data Charter and Technical Appendix, June 2013**
- Science International Accord on Open Data in a Big Data World, Dec 2015: http://bit.ly/opendata-bigdata
- EC H2020 Open Data Policy Pilot, 2014; Adoption of FAIR Data Principles, 2017.
- China State Council Law on Management of Research Data, March 2018
- Expert Group Report 'Turning FAIR into Reality' Nov 2018 https://doi.org/10.2777/1524

# Issues for Data and Diplomacy

- Data Diplomacy as a subset of science diplomacy.

- A lot of significant areas of science diplomacy have important data components.

- **Data is an important subject for diplomacy; data - and the evidence derived from it - is an important input to diplomacy.**

- There is a need to coordinate activities around data internationally.

  - agreements, coordination and collaboration around data infrastructures, data availability, data policies, data standards.

  - need to demonstrate the national interest in these global interests.

- (Scientific) data collection and stewardship as part of international agreements and coordination: who pays for data infrastructures and collections?  How is the cost shared, how are they sustained?

- Sendai Framework, SDGs: requires further 'diplomacy' to achieve optimal alignment of data (definition of hazards, coordination of 'disaggregated' data, reporting protocols etc).

# CODATA and Data Diplomacy

1. 20 year review of GBIF - global biodiversity information facility

2. How do we coordinate data for the Sendai framework?

3. CODATA-University of Arizona Center of Excellence in Data for Society - data diplomacy will play a significant part.

   - From Spacecraft to Statecraft: The Role of Earth Observation Satellite Data in Supporting the United Nations Sustainable Development Goals

# Interdisciplinary Research and Use Beyond the Designated Community

- Major interdisciplinary research issues depend on the integration of data and information from different sources.

- Fundamental importance of agreed vocabularies and standards.

    - Fundamental to integration of social science, geospatial and other data

    - Essential to effective interface of science and monitoring (e.g. Sendai, SDGs, sustainable cities)

    - Recent CODATA work: LOD for Disaster Research, Nanomaterials Uniform Description System

- Huge opportunities but significant challenges.

- **Better exploitation of data resources for interdisciplinary grand challenge research areas is the epochal challenge of the 21st century.**
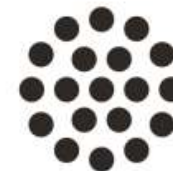
# Initiative for Data Interoperability and Integration

- **Strand 1 addresses important application domains** (**infectious disease, resilient cities, and disaster risk**): They have been chosen as major issues where relevant data exists and is accessible, where data integration is a tractable objective, and where there are existing communities of practice that are willing to collaborate.

  - ➢ **Demonstrate the tractable, achievable benefits of data interoperability.**

- **Strand 2 provides data science support for the pilots** and by extension other disciplines of science that have not yet developed the standards (vocabularies, ontologies, etc) that are necessary for effective data integration. … Formalisation of the discipline-specific vocabularies is an essential pre-requisite for integration of data from different disciplines.

  - ➢ **Drawing generic lessons on the actions and support required to promote interoperability and data integration.**

# Data Integration Initiative: Disaster Risk Reduction Pilot

- Working with **members of CODATA TG on Linked Open Data for Global Disaster Risk Research, Public Health England**

- Strong links with ISC Programme on Integrated Research on Disaster Risk

- Pilot focuses on challenges of reporting mortality in the context of the **Sendai Framework for Disaster Risk Reduction**.

- Complementary to the other pilots: significant policy dimensions, looking at different levels of data.

**Sendai Framework for Disaster Risk Reduction 2015 - 2030**

1. Aggregated data for reporting < > 'disaggregated data' < > data sources.

2. Significant definitional challenges, particularly in relation to cause of death, dead or missing;

3. Article from Dagstuhl Workshop on possible interoperability solutions for Sendai data.

**International Science Council**
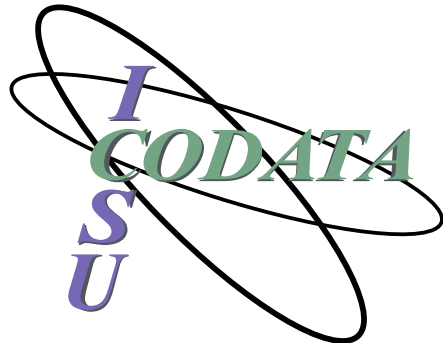
# The CODATA-UA
# Center of Excellence in
# Data for Society



Center
of
Excellence
in
Data for Society

www.ceds.arizona.edu
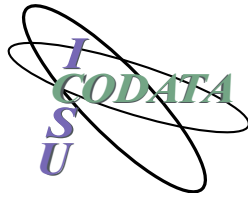
The Committee on Data of the International Science Council(CODATA) and the University of Arizona (UA) are pleased to announce their new partnership in the Center of Excellence in Data for Society (CEDS). CEDS studies the practices, public policies, and uses related to data in all its forms in the U.S. and across the globe.
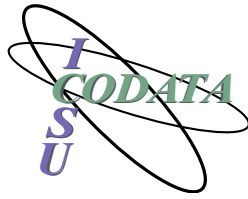
www.ceds.arizona.edu

# From Spacecraft to Statecraft:

## The Role of Earth Observation Satellite Data in Supporting the United Nations Sustainable Development Goals

This innovative project will use Earth observation satellite data to help nations achieve selected SDGs. We are seeking a nation to act as a case study to exemplify the collaborations, tools, and structures that result in effective data diplomacy leadership.

www.ceds.arizona.edu

# Questions for Data and Diplomacy

- Actions designed to directly advance a country's national needs
  - Data and national science policy (data security, data gathering exercises, national science capacity and investment)
- Actions designed to address cross-border interests
  - Particularly important in terms of data gathering in relation to (soft or hard) treaty obligations (hydrology, hazards, biodiversity, pollution)
- Actions primarily designed to meet global needs and challenges
  - Global coordination, avoiding the prisoners' dilemma, no loss of competiveness
  - Shared or coordinated investments, infrastructure, standards
  - Data for shared global challenges (environment, biodiversity, disasters etc)
  - Global responsibilities and equity (helicopter research, OA)
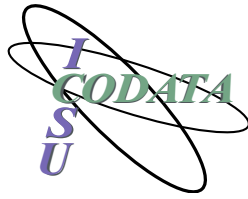
# PARTNERSHIP OPPORTUNITIES

CEDS seeks partners to help create a vision and build a network that will establish a strong data diplomacy infrastructure for users across sectors and capacities throughout the globe.

During our pilot phase (through June 2020), we are accepting nominations for Chairs on our Governing Board, Heads of our Program Offices, Endowed Programs, and Sponsored Projects.

www.ceds.arizona.edu
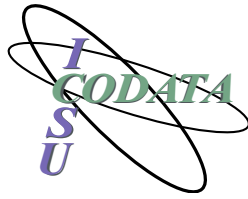
# RESEARCH OPPORTUNITIES

CEDS welcomes researchers who want to study

- data diplomacy,

- big data practices,

- data laws and regulations,

- data sharing and protection across sectors,

- citizen data, and

- other issues in the storage, use, management, and needs of data researchers and users.

www.ceds.arizona.edu

# Thank you for your attention!

Dagstuhl Workshop Slide Credits: Joachim Wackerow, Steve McEachern and Simon Cox

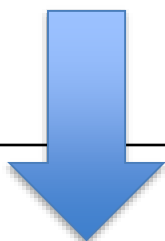## Simon Hodson
## Executive Director CODATA
www.codata.org
http://lists.codata.org/mailman/listinfo/codata-international_lists.codata.org
Email: simon@codata.org
Twitter: @simonhodson99
Tel (Office): +33 1 45 25 04 96 | Tel (Cell): +33 6 86 30 42 59

CODATA (ICSU Committee on Data for Science and Technology), 5 rue Auguste Vacquerie, 75016 Paris,
FRANCE

**Data for/in Diplomacy**
(What data is needed in relation to diplomatic agreements and how is it used in practice?)

**Diplomacy for Data**
(What are the necessary agreements, structures, mechanisms, organisations etc. needed to allow optimum access to data needed in relation to the SDGs?)
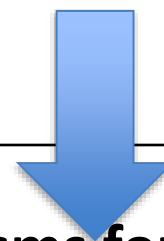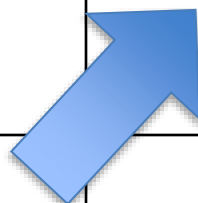
**Data to Evidence for the SDGs**
(What EO data are needed to generate evidence for decision-making in relation to the SDGs?)

**Mechanisms for Data and Evidence for the SDGs**
(What are the agreements, structures, mechanisms, methodologies needed to use data to create evidence for informed decisonmaking?)

**Toolkit for Data Diplomacy (for EO Data, Evidence and the SDGs)**

# Questions for Data and Diplomacy

**Data and Diplomacy**

- There is a need to coordinate activities around data internationally.

  - agreements, coordination and collaboration around data infrastructures, data availability, data policies, data standards.

  - need to demonstrate the national interest in these global interests.

- (Scientific) data collection and stewardship as part of international agreements and coordination: who pays for data infrastructures and collections?  How is the cost shared, how are they sustained?

- Sendai Framework: requires further 'diplomacy' to achieve optimal alignment of Sendai data (definition of hazards, coordination of 'disaggregated' data, reporting protocols etc).

# Questions for Data and Diplomacy

**Data, Science Policy and Diplomacy**

- Open Access (e.g. Plan S) can be construed as a topic for science policy and science diplomacy.

- So are data sharing policies, licensing regimes etc.

- Proliferation and alignment of data policies?

- Data policies in the African Grant Funding Councils initiative?  Data policies and coordinated funding programmes?

- North-south disparities in data infrastructure; helicopter research.

# EO Data and
# Science/Data Diplomacy Typology

- Actions designed to directly advance a country's national needs
  - ➤ Access and use of EO Data and national interests
- Actions designed to address cross-border interests
  - ➤ EO data and e.g. hydrology, hazards, biodiversity, pollution
- Actions primarily designed to meet global needs and challenges
  - ➤ Shared or coordinated investments, infrastructure, standards
  - ➤ Data for shared global challenges (environment, biodiversity, disasters etc)
  - ➤ SDGs 2 (food security), 11 (sustainable cities), 14 (sustainable oceans), 15 (biodiversity)