

Reprise on the COINAtlantic Data Accessibility Self- Assessment Tool: Adding FAIRness.

Andrew Sherin, Alexi Baccardax Westcott
COINAtlantic Secretariat



Office of the Auditor General of Canada

“Solid, objective, and accessible information is essential to identify and respond to the quickening pace and complexity of environmental change, in Canada and globally.”

2010 Fall Report of the Commissioner of the Environment and Sustainable Development

A Brief history of Research Data Management

OECD



- 2007 OECD: Principles and Guidelines for Access to Research Data from Public Funding
- 2015 Portage: Shared Stewardship of Research Data
- 2016 FAIR: Guidelines on FAIR Data Management in Horizon 2020
- 2018 Canadian Tri-council draft policies on:
 - Institutional strategy to support researchers
 - Researcher requirements for DMPs, Data Deposit
- 2014 Center for Open Science: Transparency and Openness Promotion (TOP) Guidelines



The FAIR Guiding Principles for scientific data management and stewardship: Overview

- TO BE FINDABLE:
 - F1. (meta)data are assigned a globally unique and eternally persistent identifier.
 - F2. data are described with rich metadata.
 - F3. (meta)data are registered or indexed in a searchable resource.
 - F4. metadata specify the data identifier.
- TO BE ACCESSIBLE:
 - A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
 - A1.1 the protocol is open, free, and universally implementable.
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
 - A2 metadata are accessible, even when the data are no longer available.
- TO BE INTEROPERABLE:
 - I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
 - I2. (meta)data use vocabularies that follow FAIR principles.
 - I3. (meta)data include qualified references to other (meta)data.
- TO BE RE-USABLE:
 - R1. meta(data) have a plurality of accurate and relevant attributes.
 - R1.1. (meta)data are released with a clear and accessible data usage license.
 - R1.2. (meta)data are associated with their provenance.
 - R1.3. (meta)data meet domain-relevant community standards.

CDAST Principles

- 1. Open Data by Default
- 2. Completeness
- 3. Primacy **replaced by Standard identification**
- 4. Timeliness
- 5. Ease of Physical and Electronic Access **now includes interoperability**
- 6. Non-discrimination
- 7. Licencing
- 8. Permanence **now recognizes certified repositories**
- 9. Usage Costs
- 10. Supporting Use
- 11. Evaluation

Selected CDAST Definitions

Open Data: the idea that data should be freely available for everyone to access, use and republish as they wish, published without restrictions from copyright, patents or other mechanisms of control. *Source: Open Government Partnership*

Interoperability: the ability of two or more systems or components to exchange information and to use the information that has been exchanged. *Source: IEEE Glossary 1991*

Certified Repository: a trusted digital repository that has undergone a recognized audit and certification process. *Source: The Center for Research Libraries 2007*

3. Standard identification mechanism

The data providing organization makes data identifiable and locatable using a standard identification mechanism.

(e.g. Digital Object Identifier (DOI))

<https://www.doi.org/index.html>

	A unique standard identification is generated for each data set and is easily locatable on the internet using the identifier.		A unique standard identification is generated for some data sets and is easily locatable on the internet using the identifier.		A unique standard identification is generated for few data sets or is not is easily locatable on the internet using the identifier.	No standard identification is generated.	n/a
	5	4	3	2	1	0	
✓							

5. Ease of Physical and Electronic Access

- 5a. Internet Access: The data providing organization (DPO) releases datasets using the internet.
- 5b. Standard Formats: The DPO releases datasets in open, freely available formats that conform to widely accepted standards and easily useable in widely used software.
- 5c. Standard Services: The DPO supports web-based services that conform to widely accepted standards.

5. Ease of Physical and Electronic Access

- 5d. Interoperability: The DPO provides policies, processes and services that maximize the likelihood of the reuse of data by other organizations / researchers and the integration of the data with other data for analytical purposes.

	Internationally approved discipline specific content standards and vocabularies are used including information on appropriate uses of the data and compliance with available (open) software applications.		Content standards and controlled vocabularies are used and well described including information on appropriate uses of the data.		The importance of content standards and controlled vocabularies is recognized and under development.	No content standards or controlled vocabularies are used.	n/a
	5	4	3	2	1	0	
✓							

How to use the CDAST

- Use repetitively to measure improvement in the effectiveness and ease of accessing data and information by users.
- The aggregate score obtained from all of the tool's ladders is less important than the individual scores for each ladder.
- Use to separately score different business units or data dissemination infrastructures within an organization.
- Organizations may want to weight their scores based upon internal priorities.
- Go to <https://coinatlantic.tools> for the pdf and on-line versions of CDAST

Future Work

- Apply CDAST to all projects where COINAtlantic Secretariat provides data management advice and services.
- Continue to promote the tool to data providing organizations.
- Encourage comments and critique of the tool.

Acknowledgements: Funding: Natural Resources Canada.

Collaborators: Dr. Bertrum MacDonald, School of Information Management, and Robert Branton, Emeritus, Ocean Tracking Network, both Dalhousie University.

The following organizations have provided comments to improve the CDAST:

Northeast Coastal and Ocean Data Partnership, NRCAN Adaptation Platform , NOAA Data Management Integration Team, ICAN 7 workshop / CoastGIS 2015, CODATA , EIUI Research group, School of Information Management, Dalhousie University

Canadian Integrated Ocean Observation System (CIOOS)



- **November 2015:** Expert Forum on Ocean Data Management
- **November 2017:** Completion of three Investigative Evaluations:
 - essential ocean variables,
 - cyberinfrastructure and
 - visualization
- **June 2018:** A RFP for the establishment of three pilot Regional Associations and CIOOS national web presence.
- **? 2018:** Agreements in place to establish the pilot Regional Associations.