A Practical Self-Assessment Tool for use by Data Providing Agencies to Monitor Data Accessibility Practices

Andrew G. Sherin[1], Alexi Baccardax Westcott[1], Robert Branton[2] & Bertrum H. MacDonald[3]

[1] Atlantic Coastal Zone Information Steering Committee Secretariat, Halifax, Nova Scotia, Canada

[2] Canadian National Committee for CODATA Canada

[3] School of Information Management, Dalhousie University, Halifax, Nova Scotia, Canada

The Atlantic Coastal Zone Information Steering Committee

▸ The ACZISC Mission
  ◦ To promote, facilitate and influence information management, policies and programs that enhance integrated coastal and ocean management in Atlantic Canada.

▸ From the ACZISC's Strategic Plan:
  ◦ Priority 2:
    • Collaborative sharing of data and information
  ◦ Strategy B:
    • Facilitating the commitment and investment of data providers to inter-organizational and public access to data and information.

coinatlantic.ca

Self-Assessment Tool:
11 Principles

▶ 1. Open Data by Default

▶ 2. Completeness

▶ 3. Primacy

▶ 4. Timeliness

▶ 5. Ease of Physical and Electronic Access

▶ 6. Non-discrimination

▶ 7. Licencing

▶ 8. Permanence

▶ 9. Usage Costs

▶ 10. Supporting Use

▶ 11. Evaluation

coinatlantic.ca

Self-Assessment Tool
11 Principles: Pedigree

a) Organisation for Economic Co-Operation and Development's Principles and Guidelines for Access to Research Data from Public Funding.

b) Government of Canada's Operating Principles for the Open Government site.

c) United States Memorandum on Open Data Policy.

d) G8 Charter for Open Data.

coinatlantic.ca

CODATA
I S U

The Tool

- The data accessibility self-assessment tool is for use in organizations to develop a benchmark to monitor progress in improving an organization's policies and procedures for providing effective and easy access to the data and information held by them.

- For each principle, one or more evaluation ladders have been developed with questions designed to assist the user in positioning their organization's policies and procedures on the ladder.

coinatlantic.ca

CODATA
I
S
U

How to Use the Tool

- Use repetitively over time to mark progress.
- The individual scores for each ladder are more important than the aggregate score: use the score to target areas to improve upon.
- Use separately to score multiple units within one organization.
- Have multiple persons from different perspectives (e.g. data manager, user, scientist) score within each unit.
- Fill it our in a workshop environment
- Weight scores based upon internal priorities.
- Some respondents may choose not to score some ladders (n/a).

coinatlantic.ca

Example of Ladder:
10. Supporting Use

10. Supporting Use

The data providing organization has developed policies and programs to assist users with data use and provides a point of contact to respond to complaints.

Policies, programs and easily found contacts have been developed to assist users and to receive and evaluate complaints.		A specific point of contact to answer questions and receive complaints about the use of the data is clearly defined and easily found.		General contacts for the organization are defined and easily found.	Contacts are not given or general contacts are hard to find.	n/a
5	4	3	2	1	0	
✓						

5
4
3
2
1
0

ICODATA
SU

8. Permanence

A. Infrastructure

◦ The data providing organization has established formal administrative responsibility for the long-term sustainability of the infrastructure required for data access and for ensuring data are effectively preserved, managed, archived and made accessible, permanently where long term retention has been determined to be necessary.

B. Data management policies and plans.

C. Retention Protocol and Assessment.

◦ The data providing organization uses accepted retention protocols and conducts periodic cost-benefit assessments to develop and refine retention protocols.
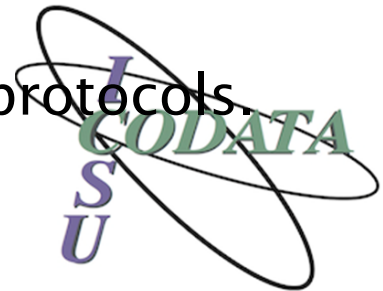
11. Evaluation

▸ The data providing organization periodically evaluates the performance of data access arrangements by **user** groups.

coinatlantic.ca

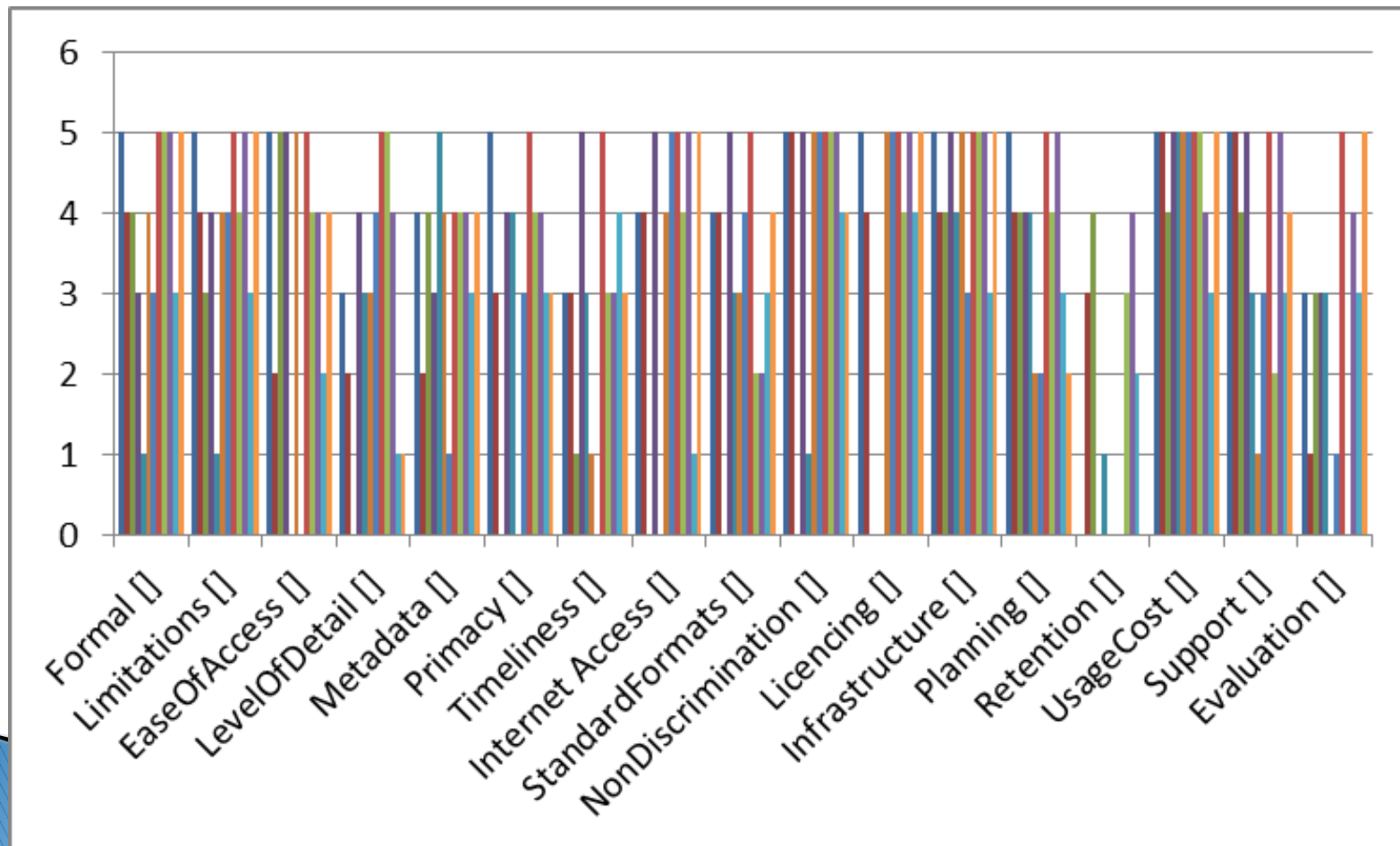# Testing the Tool So Far

- Poster at Coastal Zone Canada 2014, Halifax
- Webinars
  - NOAA Data Management Integration Team
  - NECODP
  - NRCan Adaptation Platform
- Presentation at ACZISC Meetings
- Meeting with Gov't of Canada geospatial data policy staff
- Mini-workshop ICAN 7 April 2015, Cape Town
- Presentation at CoastGIS SDI workshop on April 2015, Cape Town

Testing the Tool So Far (1)
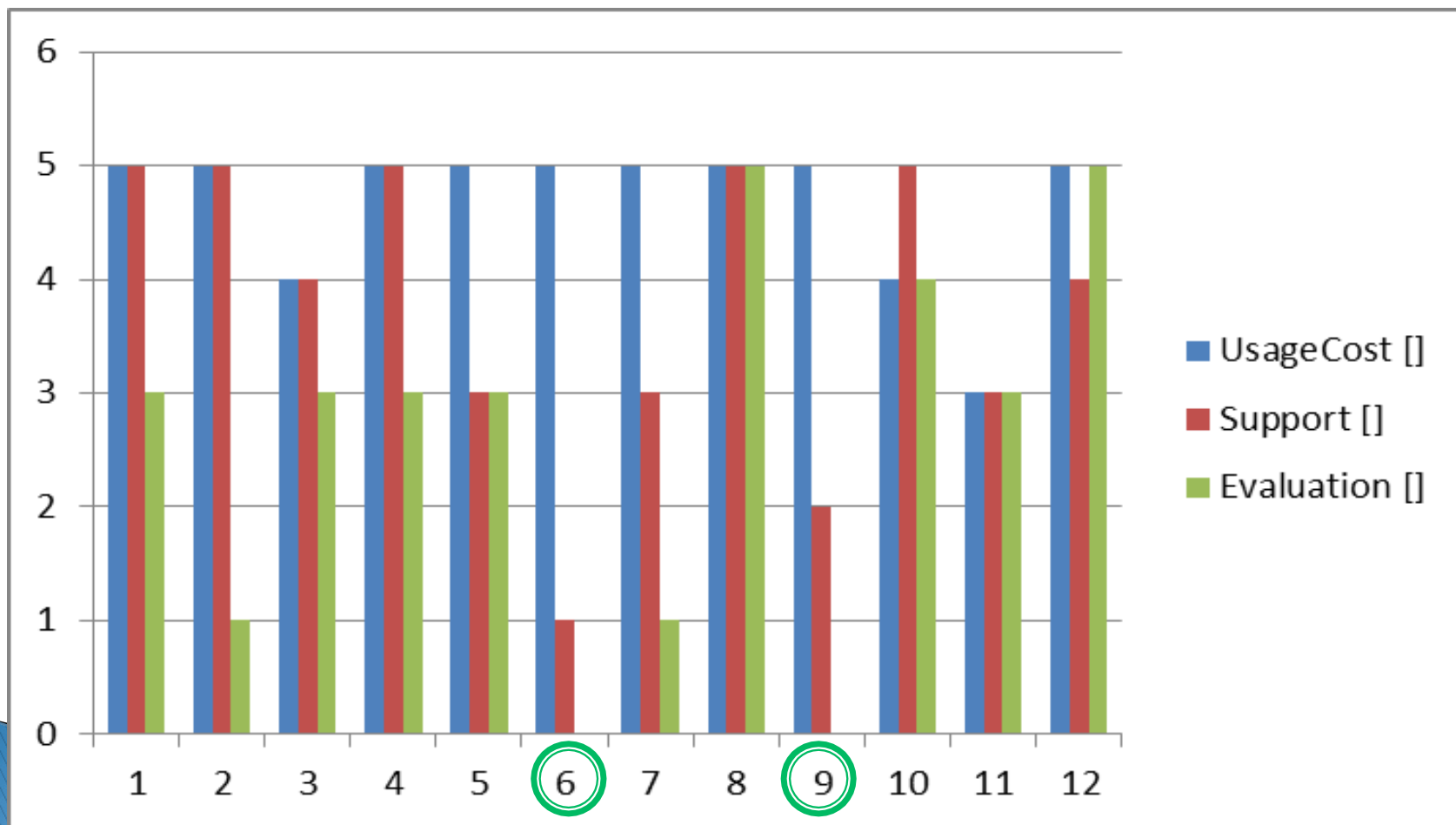
▸ Results for 12 responses / All principles

Testing the Tool So Far (2)
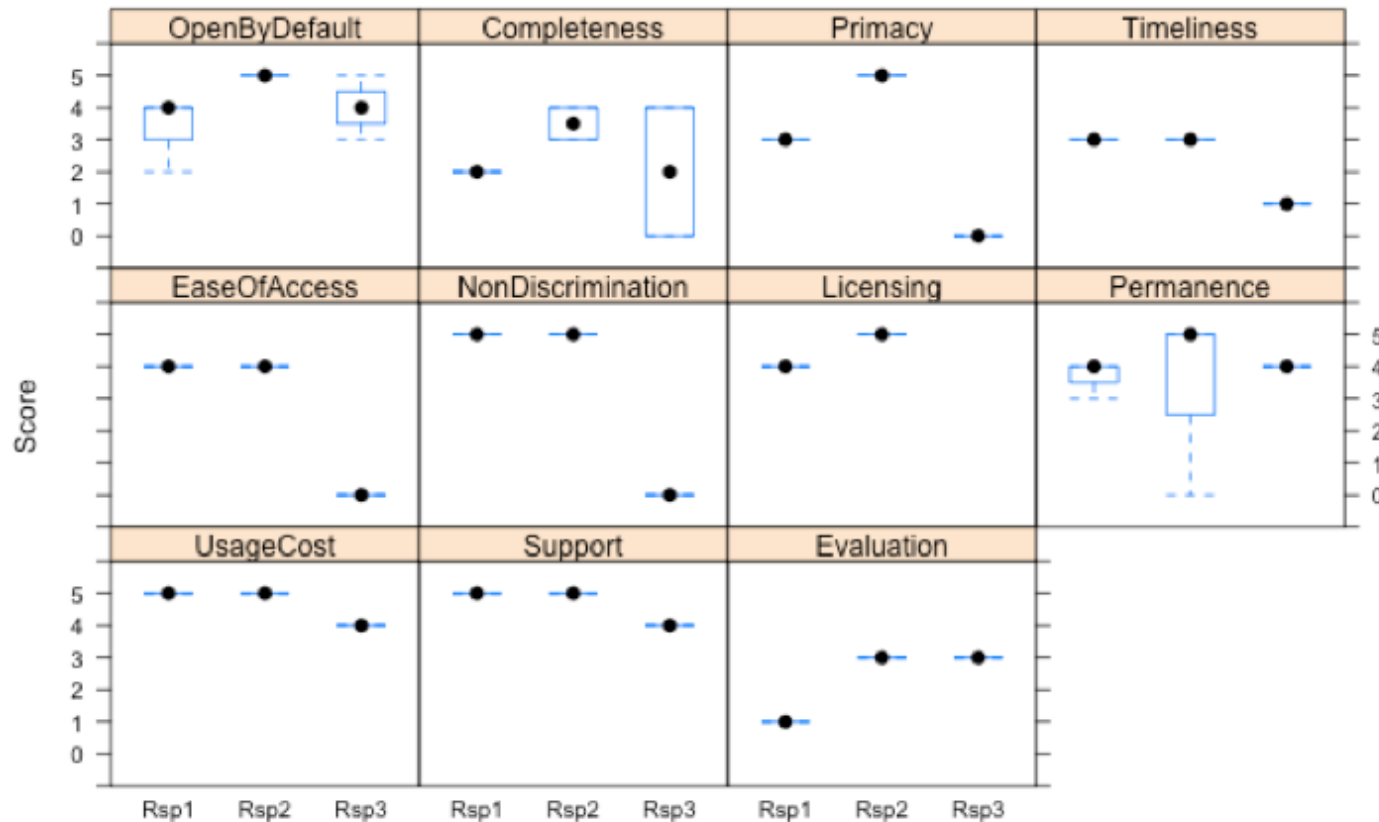
▸ Results of for the Permanence Principle
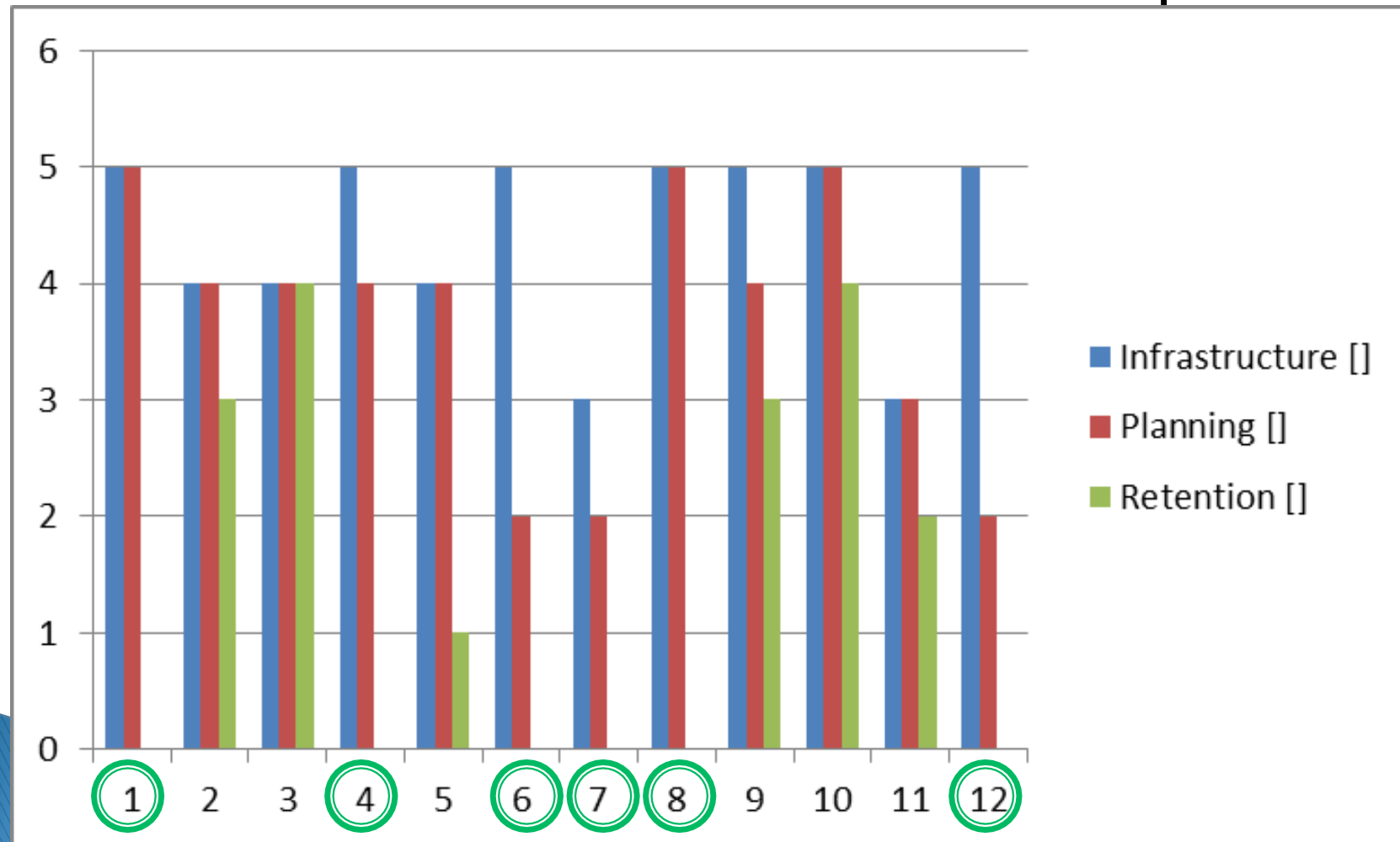
Testing the Tool So Far (3)

▸ Results for Supporting Use & Evaluation

3 Responses / 1 Organization

Accessibility Scores by Respondant and ACZISC Principal

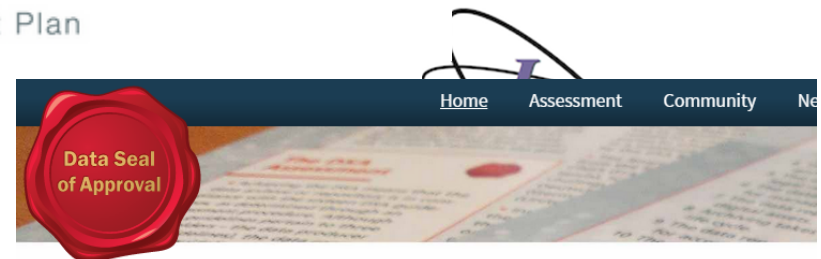# Other Assessment Initiatives

▸ Data Management Planning Tools

  ◦ https://www.dataone.org/data-management-planning

  ◦ http://www.dcc.ac.uk/dmponline

▸ Data Seal of Approval

  ◦ http://datasealofapproval.org/en/

"A goal without a plan is just a wish." Antoine de Saint-Exupery (1900 -1944)

DMPTool
Build your Data Management Plan

DMP ONLINE

coinatlantic.ca

Data Seal of Approval

Home   Assessment   Community   Ne

Towards sustainable and trusted data repositories

ACZISC / DMP Online Comparison

Results were summarized and compared using this simple table.

ACZISC - Accessibility Benchmark	DCC - DMPOnline							
	Data Collection	Data Sharing	Document-ation	Ethics	Responsi-bilities	Selection	Storage	Total
Completeness	1	.	1	.	.	.	.	2
Ease Of Access	.	2	.	.	.	.	.	2
Evaluation	.	.	.	.	1	.	.	1
Licensing	.	.	.	1	.	.	.	1
Non-Discrimination	.	1	.	.	.	.	.	1
Open By Default	.	2	.	1	.	.	.	3
Permanence	.	.	.	.	1	1	1	3
Primacy	1	.	.	.	.	.	.	1
Support	.	.	.	.	1	.	.	1
Timeliness	1	.	.	.	.	.	.	1
Usage Cost	.	.	.	.	1	.	.	1
Total	3	5	1	2	4	1	1	17

Present Status of the Tool

- Paper and PDF versions:
  - http://coinatlantic.ca/documents/ Data_Accessibility_Benchmark_Organizational_Self_Assessm ent_Tool.pdf
- Google Form Version
  - https://docs.google.com/forms/d/1MJDBaBK8y- PE1XI6YmZ4BlqTGYwPmYGACNjIDVhXXsM/edit?pli=1
- Pilot stage to test and revise the tool: Inviting data providing organizations to try it out and provide feedback.
- Interested in receiving actual scores but organizations will not be identified.

coinatlantic.ca

Questions / Discussion

- Acknowledgements
  - Financial support: Adaptation Platform Program, Climate Change Impacts and Adaptation Division of Natural Resources Canada
  - Tool development: Dr. Bertrum MacDonald, School of Information Management, Dalhousie University
 - Tool testing and Google Form development: Robert Branton, CODATA Canada
- Contact us: a.sherin@dal.ca (902) 494-6184
- Website: coinatlantic.ca
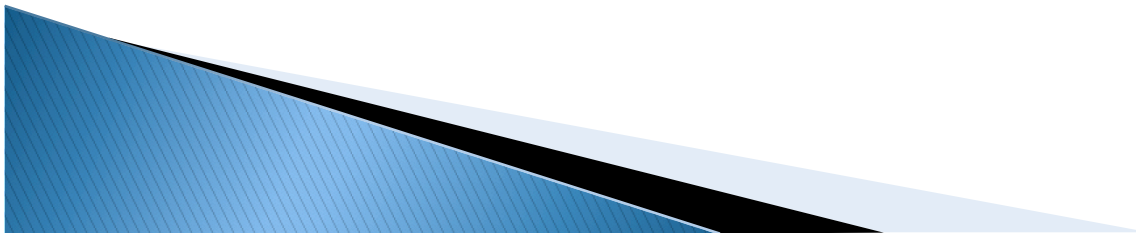
1 Open Data by Default

▸ A. Formal Policies and Limitations
  ◦ The data providing organization has adopted policies which presume open data.
  ◦ Score 5: Policies presume that data is open and limitations are clearly defined.

▸ B. Ease of access to information on policies
  ◦ Information is disseminated online regarding the data providing organization's data access policies, etc.
  ◦ Score 5: Policies, etc. are prominently displayed on websites and are easily found.
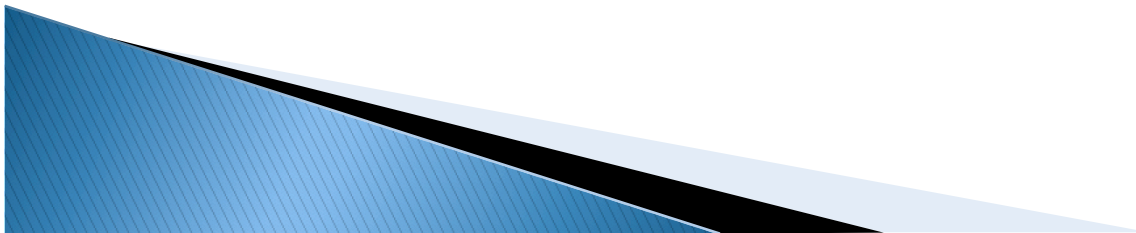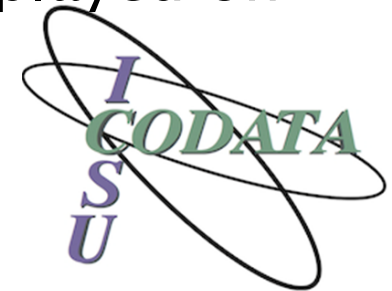
2. Completeness

A. Level of Detail, Data
- The data providing organization releases datasets that are as complete as possible at the finest level of granularity available.
- Score 5: Datasets are complete and detailed.

B. Metadata
- The data providing organization releases metadata that defines and explains the data, explains how derived data was calculated, and states the standards employed for the data, metadata and quality assessment.
- Score 5: Complete and readily available metadata using an acceptable standard for every data set.

3. Primacy

- The data providing organization releases primary source data; this includes the original raw data collected and metadata on how the data was collected.
- Score 5: Raw / primary data is released with collection metadata.
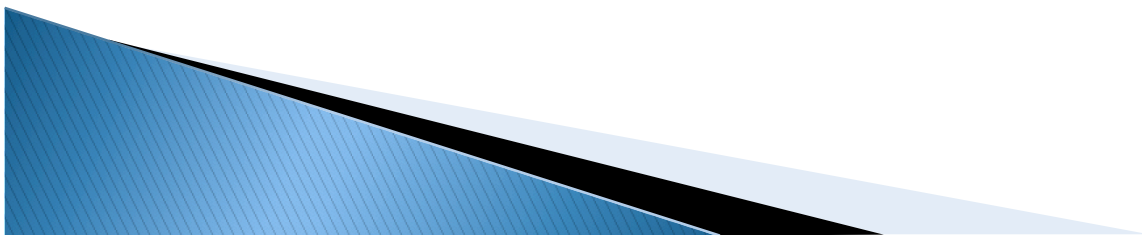
4. Timeliness

▸ The data providing organization releases datasets to the public in a timely fashion to maintain the value of the data, with priority given to data whose utility is time sensitive.

▸ Score 5: Policy states, and in practise, datasets are released in a timeframe that maintains the full value of the data.
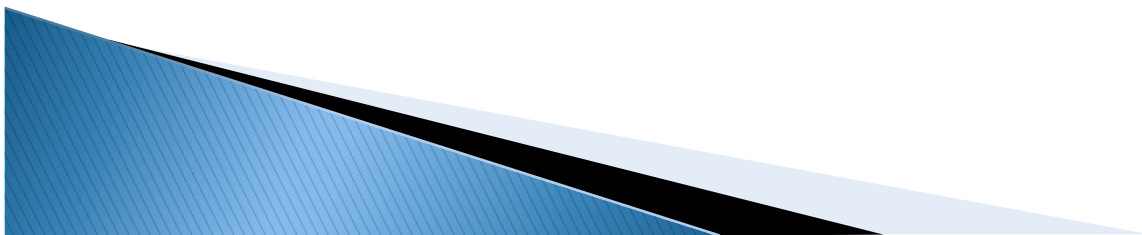
5. Ease of Physical and Electronic Access

A. Internet Access

- The data providing organization releases datasets using the internet.
- Score 5: Datasets are easily accessible on the internet.

B. Standard Formats

- The data providing organization releases datasets in open, freely available formats and web-based services that conform to widely accepted standards.
- Score 5: Datasets are accessible as standard internet services, e.g., OGC for spatial data.
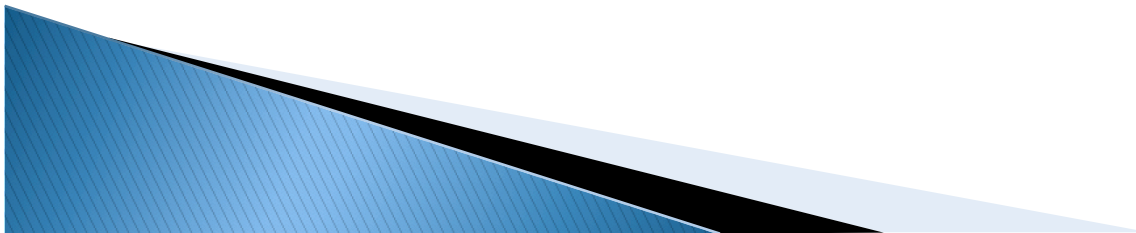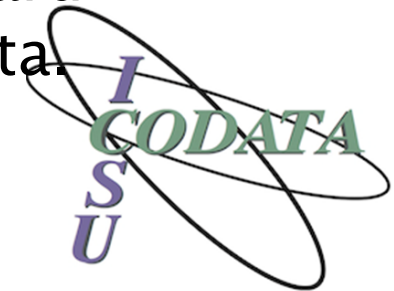
6. Non-discrimination

▸ The data providing organization provides barrier free and non-discriminatory access to data at any time without having to provide identification or justification for access.

▸ Score 5: Datasets are accessible to everyone without disclosing reason for use or identity.
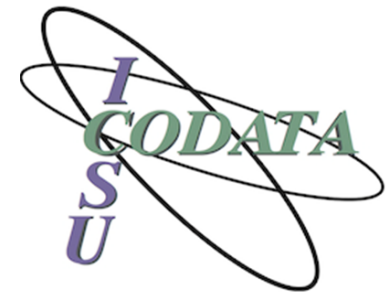
7. Licencing

- The data providing organization releases datasets under an open licence with minimal restrictions pertaining only to intellectual property, personally-identifiable and sensitive information.
- Score 5: An open license consistent with constraints permitted by Creative Commons Attribution (CC BY) license.
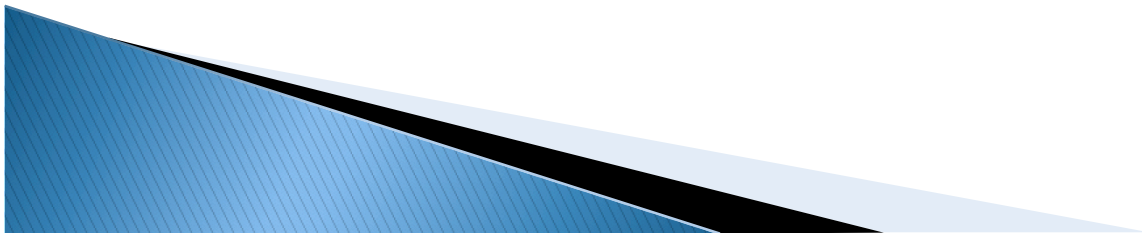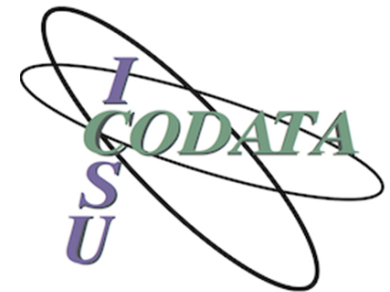
9. Usage Costs

- The data providing organization releases datasets free of charge.
- Score 5: Datasets are provided free of charge.