Metadata Training

John Clark

NSERC ResNet
2021-10-29

What is Metadata?

- "Data about data"
- Critical for data discovery and reuse
- Structured document of elements and attributes
  - Description
  - Keywords from shared thesaurus
  - Sources and license information
  - Processing steps/methods
  - Attribute definitions
  - Persistent identifier (DOI)
  - Contact information
- Standards and specifications defined by a common schema

What is Metadata?

- "Data about data"
- Critical for data discovery and reuse
- Structured document of elements and attributes
  - Description
  - Keywords from shared thesaurus
  - Sources and license information
  - Processing steps/methods
  - Attribute definitions
  - Persistent identifier (DOI)
  - Contact information
- Standards and specifications defined by a common schema

Why do we need it?

- ● For the network
  - ○ Collaboration is the heart of our network
  - ○ Make connections across Landscapes and Themes
  - ○ Shared repository of knowledge
  - ○ Standardized for interoperability
- ● For everyone
  - ○ Reproducibility, transparency
  - ○ Reduce duplication of effort
  - ○ Public interest
- ● For you
  - ○ Increasingly required by funder/publisher
  - ○ Enhance your project management
  - ○ Share data for broader use
  - ○ Incorporate into data driven apps
  - ○ Amplify the impact of your work
    - ■ Published data is citable data

Piwowar et al. 2007 doi:10.1371/journal.pone.0000308
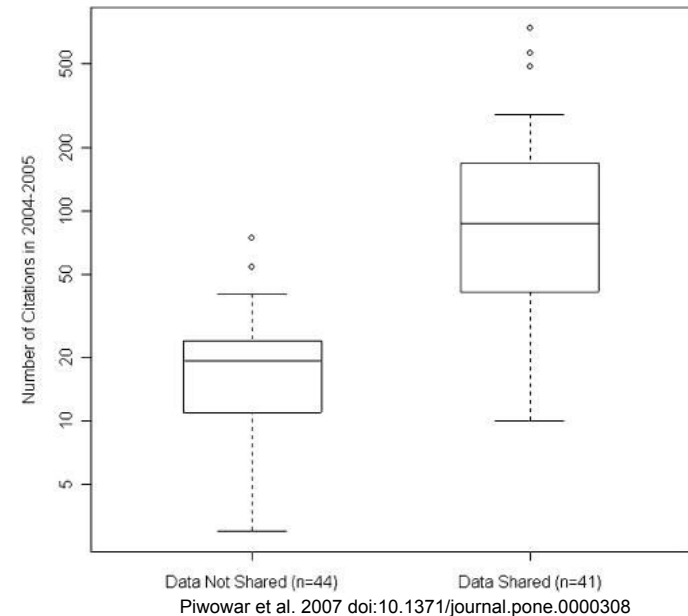
Why do we need it?

- FAIR Data Principles
  - **F**indable
    - Metadata and data should be easy to find for both humans and computers
  - **A**ccessible
    - Must be able to retrieve after finding
  - **I**nteroperable
    - Data need to interoperate with applications or workflows for analysis, storage, and processing
  - **R**eusable
    - Optimise the reuse of data. Metadata and data should be well-described so that they can be replicated and/or combined in different settings

Feedback: Poll
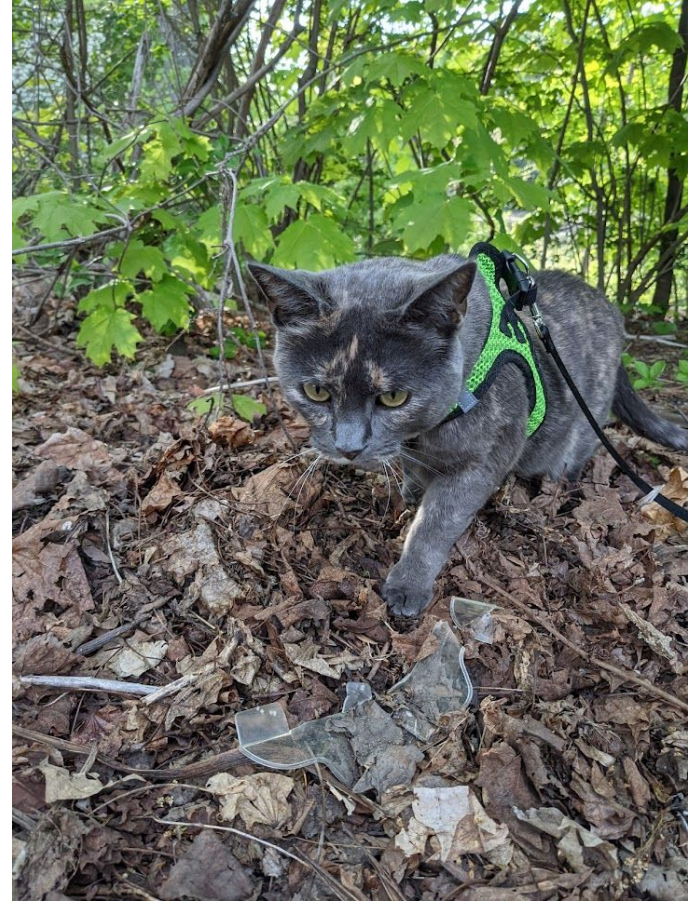
https://bit.ly/MetadataPoll

Building Blocks

- Structured data document
  - Element, attribute pairs
  - Classes and relationships
- Machine readable/interoperable
  - Date/extent standard encoding
  - Keywords from common ontology
- Output/storage can be various file types
  - JSON, XML
- Typically included in data catalogue
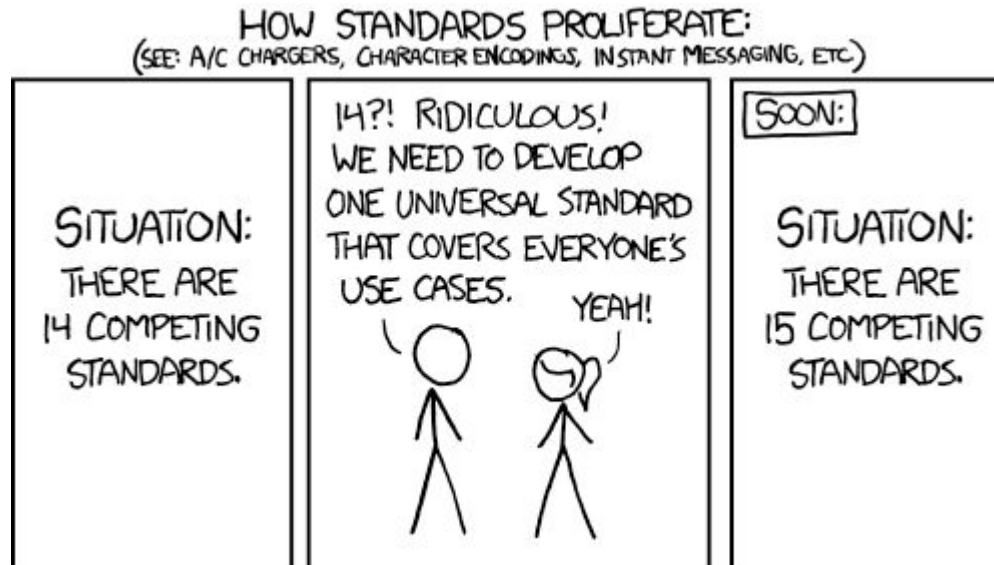 - CSW

In the wild

ResNet Metadata: Target Standard

- ISO 19115 Harmonized North American Profile
- Developed in 2000s
- Replaces Content Standard for Digital Geospatial Metadata (CSDGM)
- Modular/flexible emphasis
- Entity relationships
- Adopted by federal government

https://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912

Metadata Challenges

https://xkcd.com/927/

Metadata Challenges

- Early adoption = multiple implementations
    - Specifications, format, validation methods vary
    - Interoperability still a work in progress
- Broad scope and modularity = complexity
- Time of contributors is a precious resource
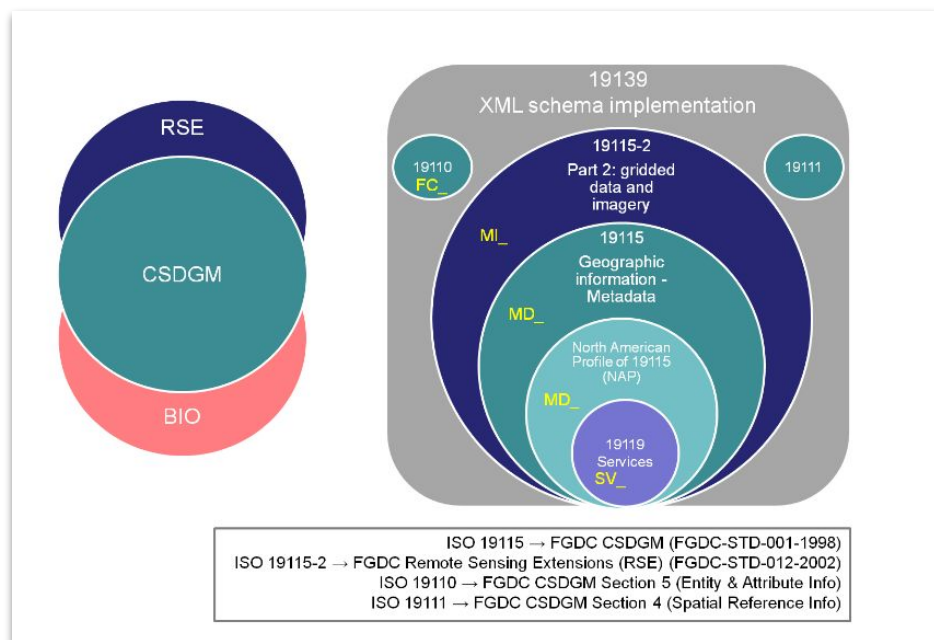
348 p

Metadata Authoring Tools

Ideal Features

- Cross platform
- Open source
- Schema Aware
- Offline/portable
- User friendly interface
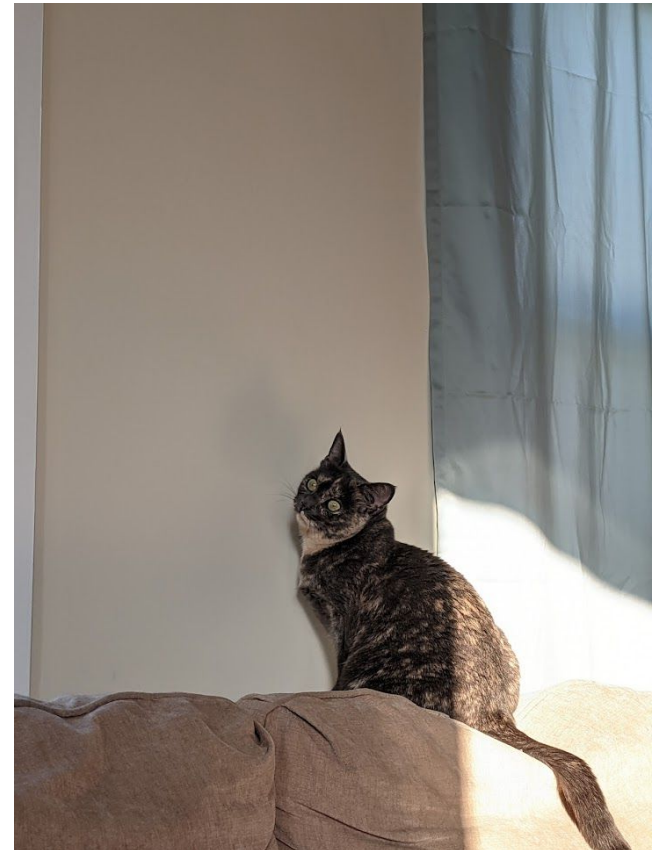
Metadata Tools: Current Landscape

Oops! It looks like what you were searching for isn't the only thing that is lost.

We are terribly sorry and hope that these links below will help find your way.

Metadata Tools: Survey

- ArcMap/ArcGIS Pro
  - Proprietary
  - Windows only
- QGIS
  - ISO/NAP not implemented
  - Plugin available but has issues
- GeoNode (Data Portal)
  - Open source
  - Web based
  - Workable UI, not perfect
- GeoNetwork
  - Similar to GeoNode
- mdEditor.org
  - In browser only, 3rd party hosted
  - Format locked to mdJSON
- (r)geometa
  - R and python packages for working with CSW metadata
  - Powerful, not user friendly

Demo

Taking a step back...

- Implementing metadata standards is an ongoing effort, at all levels
- Typical user should not have to handle underlying complexities
- Formatting minutiae are important, but the content is critical

- Authoring that underlying content is something you do already!
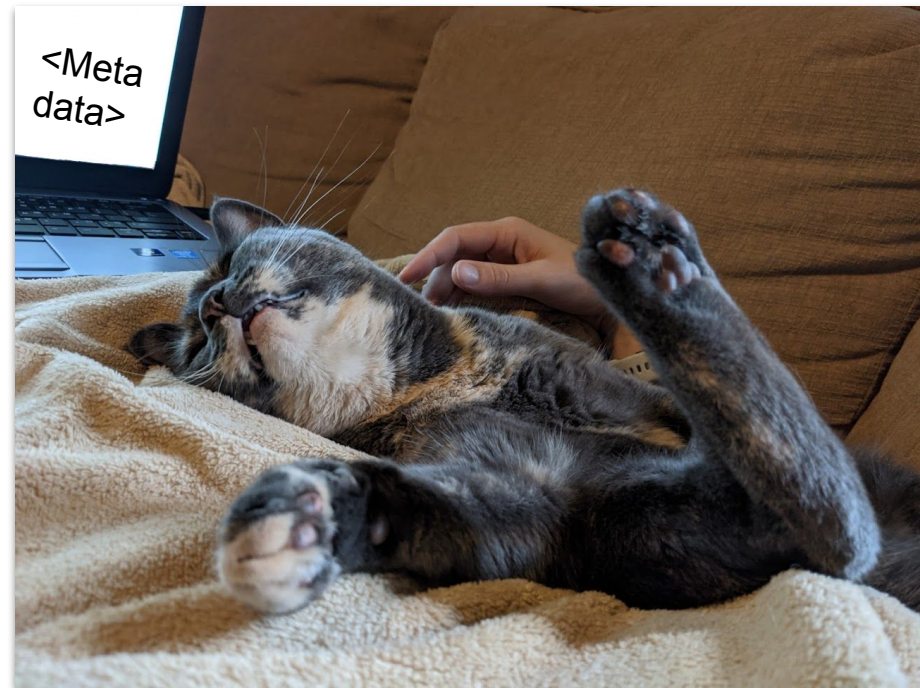
Metadata Questionnaire

20 human-friendly questions covering the core requirements.

https://bit.ly/MetadataQuest

Metadata Questionnaire: Feedback

- What questions will be challenging to answer?
- Are there any gaps/aspects of your work that aren't addressed?

https://bit.ly/MetaFeedback

What's Next?

Feedback

- https://bit.ly/MetadataPoll
- https://bit.ly/MetaFeedback

Resources

- Metadata Questionnaire:
  https://bit.ly/MetadataQuest
- Engage on Slack (#data)
- Reach out directly

Workshops

- Data Portal Orientation:
  Thursday, November 18 1-2pm EST

Ongoing

- Data Policy
- User Guides

Data Office Hours

Everything data: discovery, processing, analysis, visualization, development, publication, puns…

All welcome!

Wednesdays 2-3pm ET

"Knock" on my door via Slack/Email
and we'll set up a Zoom session

Contact

John Clark
NSERC ResNet
Data Manager, Dashboard Developer
john.clark3@affiliate.mcgill.ca

Metadata Tools: Survey

- ArcMap
  - https://desktop.arcgis.com/en/arcmap/latest/manage-data/metadata/creating-standard-compliant-metadata.htm
- ArcGIS Pro
  - https://pro.arcgis.com/en/pro-app/latest/help/metadata/create-nap-metadata.htm
- GeoNode (Data Portal)
  - https://docs.geonode.org/en/master/usage/managing_layers/layer_metadata.html