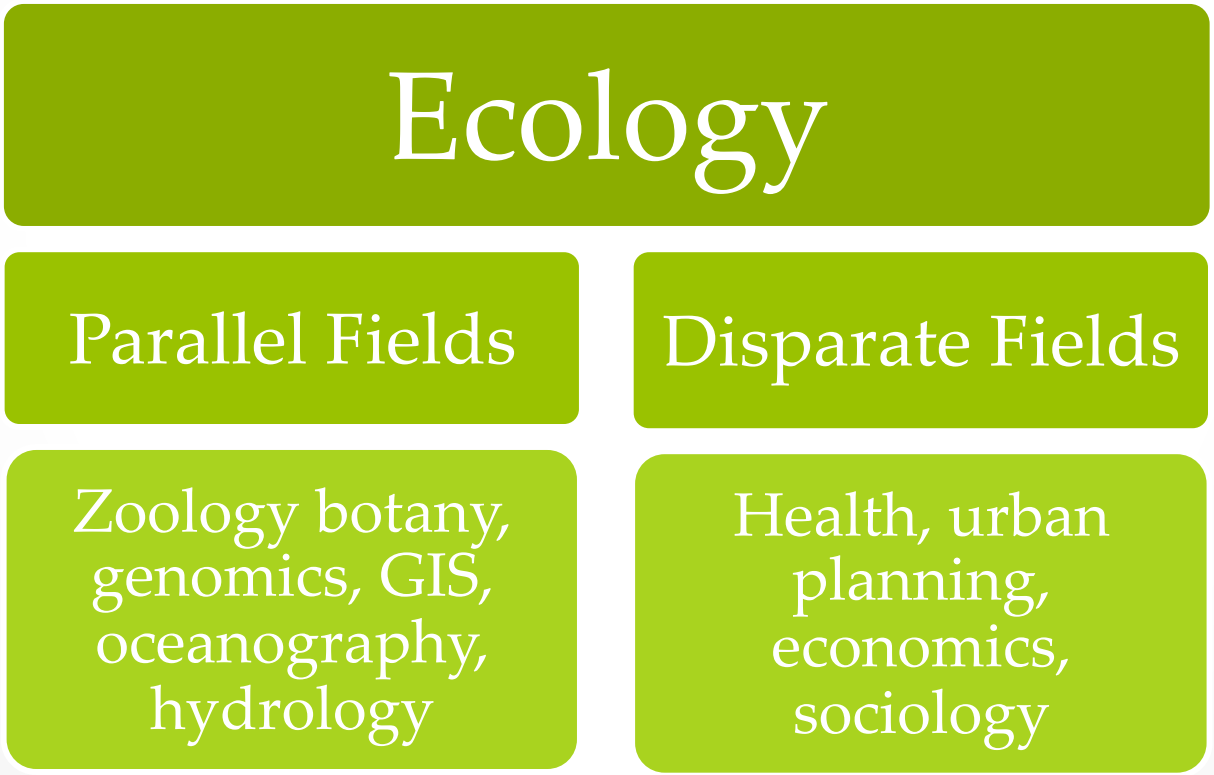
Meaningful Metadata
and the Enhancement of Data Sharing

Christine Adams
cbadams@ualberta.ca
School of Library and Information Studies
University of Alberta
FIP 2016

Ecology 101

- The study of the relationships in natural communities, typically from a biological standpoint



Callenbach, Ernest. (2008). Ecology: A pocket guide. Berkley, University of California Press.
Reichman, O.J., Jones, M.B, Schildhauer, M.P. (2011). Challenges and Opportunities of Open Data in Ecology. *Science*, 331, 703-705.

Sharing is Caring

- Benefits to the discipline
  - Prevents redundant data collection
  - Efficient use of resources
  - Increases the transparency of science
  - Advancement of discovery and public good

- Benefits to the researcher
  - Stipulation for publication or funding
  - Increases citation rates

Piwowar, H.A. (2013). Almetrics: Value All Research Products. *Nature*, 493(7431), 159.

Data Should Be…

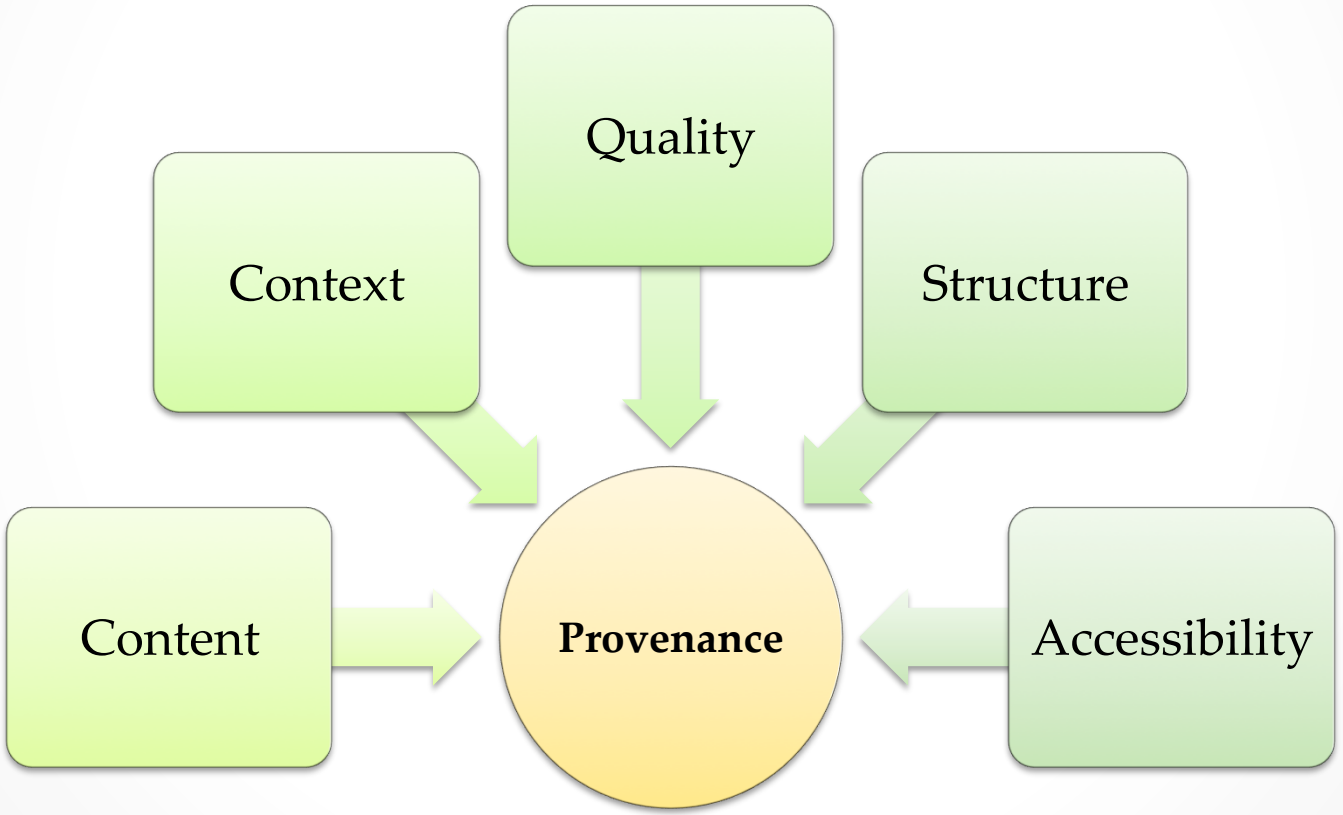
- ***Discoverable*:** Data must be capable of being located, identified, and generally assessed through simple tools available to many communities.
- **Open:** Data should generally be openly accessible.
- **Linked:** Data should be interrelated and connected.
- **Useful:** Data must be able to be used for a practical, advantageous purpose or in several ways by defined but possibly very different users.
- **Safe:** Data should be protected from risk, corruption, and loss; now and over the long term

(Parsons *et al.*, 2011)

Parsons, M.A., Godøy, Ø., LeDrew, E., deBruin, T.F., Danis, B., Tomlinson, S., Carlson, D. (2011). A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science. *Journal of Information Science*, 37(6), 555-569.

Metadata

Data Provenance



- Michener, W. (2015). Ecological Data Sharing. *Ecological Informatics*, 29, 33-44.

Metadata Quality

- Notoriously poor quality
  - Tools have been made in an attempt to bypass metadata
- No agreed upon standard for metadata or vocabulary
- Even in a regulated environment usage can be inconsistent
  - An examination of records for spatial data from GeoDiscover Alberta found data quality metadata elements were misused in over a third of records

Pocas, Isabell, Goncalves, Joao, Marcos, Bruno, Alonso, Joaquim, Castro, Pedro, Honrado, Joao. (2014). Evaluating the fitness for use of spatial data sets to promote quality in ecological assessment and monitoring. *International Journal of Geographical Information Science*, 28(11), 2356-2371.

Librarians to the Rescue!

Librarian as a Translator

- Help researchers understand standards and increase their willingness to share

- In the case of GeoDiscover
  - Updated standards to reflect the data creators' needs
  - Creation of supporting documentation tailored to the requirements of our user community
  - Personalized one-on-one training sessions
  - Increased engagement from our data contributors

Sayogo, D.S., Pardo, T.A. (2012). Exploring the Determinants of Scientific Data Sharing: Understanding the Motivation to Publish Research Data

Librarian as a Scientist

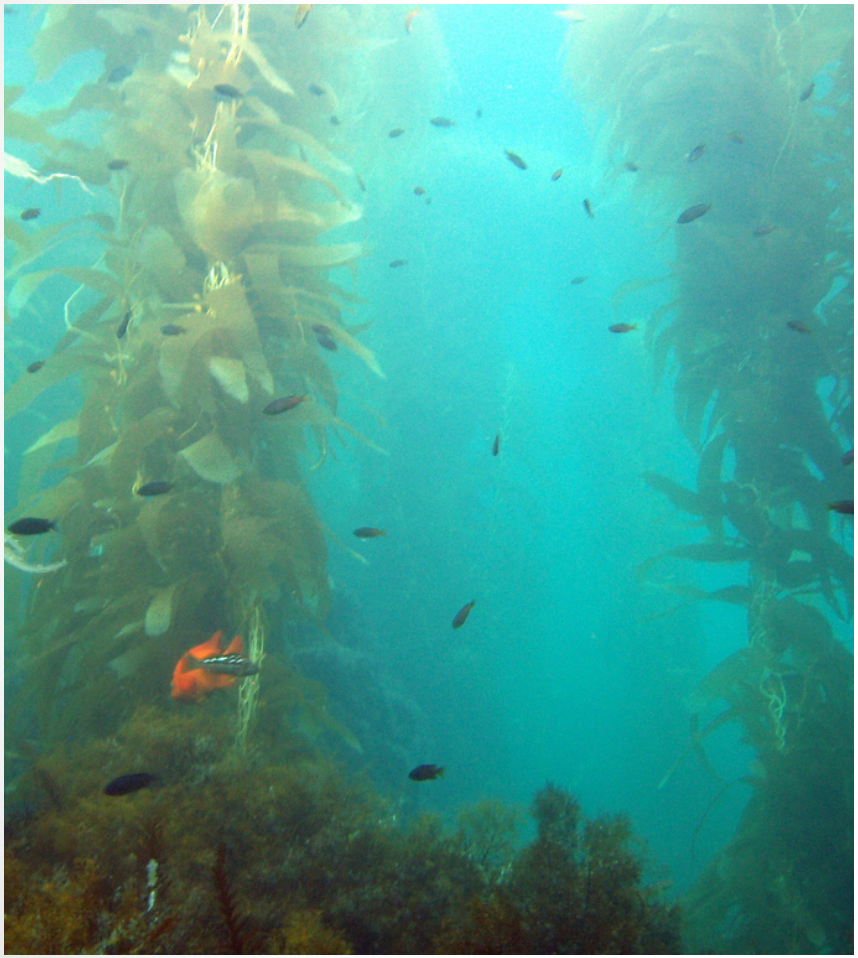
- Value of working closely with researchers
  - Improve metadata quality through increased communication
  - Increased trust between parties
  - Better understanding of the tacit subject knowledge of the researcher

- Help with the research to understand researchers' workflow

Parsons, M.A., Godøy, Ø., LeDrew, E., deBruin, T.F., Danis, B., Tomlinson, S., Carlson, D. (2011). A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science. *Journal of Information Science*, 37(6), 555-569.

Research Ecosystem

Librarian as a Keystone Species



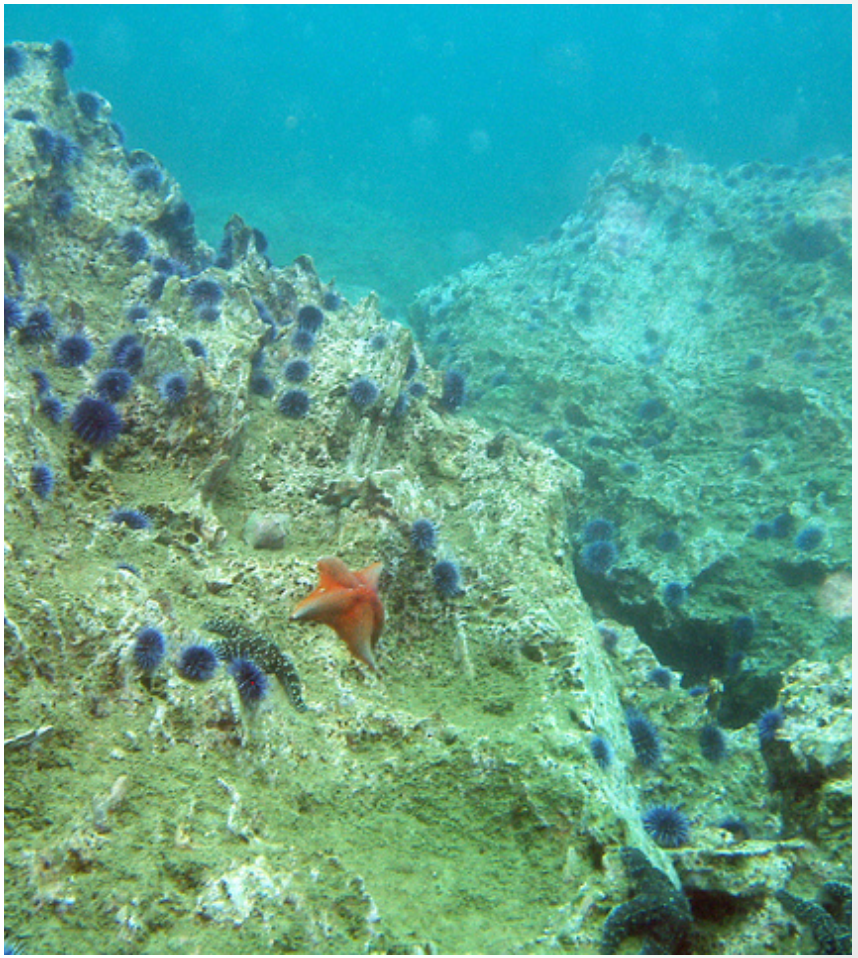


Photo Credit: Dana Roeber Murray

- Retrieved from: https://www.flickr.com/photos/healthebay/4834879645/in/album-72157624471592777/ https://www.flickr.com/photos/healthebay/4853375085

Be the Best Otter You Can Be!

Engage with researchers to understand what they need to portray in their metadata and help them find a system that improves their capacity to share data.

Photo Credit: Matt Knoth

Retrieved from: https://upload.wikimedia.org/wikipedia/commons/2/25/Sea_otter_with_sea_urchin.jpg

References

- Ashander, J., M. Krkosek, & M. Lewis, 2012. Aquaculture-induced changes to dynamics of a migratory host and specialist parasite: a case study of pink salmon and sea lice. *Theoretical Ecology. 5*, 231-252.
- Callenbach, Ernest. (2008). Ecology: A pocket guide. Berkley, University of California Press.
- Michener, W. (2015). Ecological data sharing. *Ecological Informatics*, 29, 33-44.
- Parsons, M.A., Godøy, Ø., LeDrew, E., deBruin, T.F., Danis, B., Tomlinson, S., Carlson, D. (2011). A conceptual framework for managing very diverse data for complex, interdisciplinary science. *Journal of Information Science*, 37(6), 555-569.
- Piwowar, H.A. (2013). Almetrics: Value All Research Products. *Nature*, 493(7431), 159.
- Pocas, Isabell, Goncalves, Joao, Marcos, Bruno, Alonso, Joaquim, Castro, Pedro, Honrado, Joao. (2014). Evaluating the fitness for use of spatial data sets to promote quality in ecological assessment and monitoring. *International Journal of Geographical Information Science. 28(11), 2356-2371.*
- Reichman, O.J., Jones, M.B, Schildhauer, M.P. (2011). Challenges and opportunities of open data in Ecology. *Science*, 331, 703-705.
- Sayogo, D.S., Pardo, T.A. (2012). Exploring the determinants of scientific data sharing: understanding the motivation to publish research data