

Data Stewardship & Sustainability Strategy Guide

This guide provides a brief, non-comprehensive overview of concepts and ideas around data stewardship. The glossary of terms and concepts below can be used to start a conversation with local librarians, archivists and other information professionals to help you with data stewardship and sustainability.

Link to the Presentation slides

- <https://goo.gl/zVuUTm>

Link to this handout:

- <https://goo.gl/ivKb4m>

Resources

Sunlight Foundation

- Open Data Policy Guidelines: <http://sunlightfoundation.com/opendataguidelines/>
- Local Municipal Data Initiative: <http://sunlightfoundation.com/policy/local/>

Code For America

- Library: <http://www.codeforamerica.org/library/>
- Data Formats: <http://www.codeforamerica.org/our-work/data-formats/>

Beyond Transparency Book

- <http://beyondtransparency.org>

Project Open Data

- <https://project-open-data.cio.gov>

Metadata Standards for Data Catalogs

- DCAT: <http://www.w3.org/TR/vocab-dcat/>
- Project Open Data Metadata Schema: <https://project-open-data.cio.gov/v1.1/schema/>

Controlled Vocabularies

- Geonames: <http://www.geonames.org/>
- IANA Media Types: <http://www.iana.org/assignments/media-types>

Glossary

Stewardship - the active management of resources over time. Stewardship usually brings together managed content with people, activities, policies, systems, and tools.

- http://en.wikipedia.org/wiki/Data_steward

Discoverability - the degree to which information resources are visible and easy to find. For a look at librarians thinking through issues of discoverability:

- <http://purl.umn.edu/48258>

Interoperability - The ability of different systems to use and exchange information through a shared format. For example, Hypertext Markup Language allows browsers and servers to constitute the World Wide Web.

- <http://www2.archivists.org/glossary/terms/i/interoperability>

Open Data & Open Access - Two ideas from the “Open” movement that argue data and content should be freely available to everyone to use and re-use without economic, technical or legal restrictions.

- http://en.wikipedia.org/wiki/Open_data
- http://en.wikipedia.org/wiki/Open_access

LOCKSS - “Lots of Copies Keep Stuff Safe”, coined by the LOCKSS project. The corollaries, “Lots of description keeps stuff meaningful; lots of services keep stuff useful; lots of uses keep stuff valuable” are from Preservation is Not a Place by Stephen Abrams, Patricia Cruse, John Kunze.

- LOCKSS Project: <http://www.dlib.org/dlib/june01/reich/06reich.html>
- Preservation is Not a Place: <http://www.ijdc.net/index.php/ijdc/article/view/98>

Curation categories - for an example of identifying categories of curation.

- <http://ukdataservice.ac.uk/media/398725/cd227-collectionsdevelopmentpolicy.pdf>

Format Specification - The Library of Congress publishes a *Recommended File Format Specifications* guide as a part of its digital preservation resources. The guide includes recommendations for datasets.

- LOC Specification: <http://www.loc.gov/preservation/resources/rfs/data.html>

Filename Conventions - This training module, developed by NASA, DataOne, NOAA, and the Data Conservancy, contains good guidance on file naming conventions.

- Descriptive File Names: <http://commons.esipfed.org/node/702>

Data Management Plans - A document created in advance of data collection that identifies what data will be collected, what formats will be used, how the data will be organized and described, and data will be shared and preserved in the long-term. Funders are increasingly requiring a data management plan as part of grant applications.

- <http://www.dcc.ac.uk/resources/data-management-plans>
- <https://www.lib.umn.edu/datamanagement/DMP>