

2017

Purdue Data Orientation Checklist: An Archival Approach for Data Users

Carly Dearborn
cdearbor@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_fscm




Part of the [Archival Science Commons](#)

Dearborn, Carly, "Purdue Data Orientation Checklist: An Archival Approach for Data Users" (2017). *Libraries Faculty and Staff Creative Materials*. Paper 13.

http://docs.lib.purdue.edu/lib_fscm/13

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Purdue Data Orientation Checklist: An Archival Approach for Data Users

Author	Carly Dearborn	
Contact	cdearbor@purdue.edu	
Data of Creation	October 14, 2016	
Date of last update	March 17, 2017	
Version	V 1.0	
Acknowledgement	Based on G.R.I.P (Graduate Research Information Portal) course "Data Management IV." Instructors: Carly Dearborn and Megan Sapp Nelson	
License	Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.	

Notes on the Data Orientation Checklist

The Data Orientation Checklist is intended to help a data user navigate an unfamiliar dataset, research material or digital collection. Whether joining a new research lab or downloading data from a public repository, navigating new data can be confusing and frustrating. This checklist was developed from the perspective of an archivist working with research data and borrows best practices from archival processing and research.

Archivists at Purdue University Libraries have been working with data curation initiatives since the mid-2000s. Early in the Libraries' work with managing research data, Dean of Libraries James Mullins saw the value of archivists' knowledge in areas such as raw "data" collection, sensitive or personal information management, and defining user groups. Mullins brought the University Archivist onto the Steering Committee for the Purdue University Research Repository (PURR) in 2011 and today archivists continue to collaborate with data librarians and faculty researchers on managing, curating, and preserving research data.

This checklist was created in 2016 for use in the Purdue G.R.I.P (Graduate Research Information Portal) course Data Management IV, instructed by Carly Dearborn and Megan Sapp Nelson.

Data Orientation Checklist

Data **context**:

- Identify who (individuals, group, organization) created the data.
- Identify the research question/hypothesis
- Identify the data capture, processing and analysis methodologies
 - Identify the instruments and measurements used
- Identify geographic location and time period

Where to look: Descriptive metadata, README files, Repository metadata, tags

Data **structure**:

- Notice how the files/data are organized
- Look for signs of normalization or standardization
 - File formats
 - Units of measurement
 - Disciplinary structure (if applicable)
- Determine the relationships among the data.

Where to look: File hierarchy, naming conventions, versioning information, README files, data dictionaries, codebook, column headers, lab manual and standard operating procedures,

Data **use**:

- Determine if you have permission to use the data
 - If applicable, verify you have appropriate permissions to utilize sensitive or controlled data (FERPA, HIPAA, PII, etc.)
- Determine if you have appropriate tools and knowledge to use the data
 - Verify data quality (age, version, completeness, accuracy)
 - Do the data require you to transform (format, syntax, semantics) them in any way to use?
 - Determine how you can get access (dumps, API, etc.) and in what forms (raw, summary) data are in.

Where to look: Access permissions, licensing, or copyright in README files, on repository platforms. Make note of possible intricacies in licensing agreements.