

Purdue University Purdue e-Pubs

Libraries Faculty and Staff Creative Materials

Purdue Libraries

2017

Purdue Data Reuse Checklist: An Archival Approach for Data Producers

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 Reuse Checklist: An Archival Approach for Data Producers" (2017). *Libraries Faculty and Staff Creative Materials*. Paper 14.

http://docs.lib.purdue.edu/lib_fscm/14

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 Reuse Checklist: An Archival Approach for Data Producers

Author	Carly Dearborn
Contact	cdearbor@purdue.edu
Date 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 Reuse Checklist

The Data Reuse Checklist is intended to help a data producer verify their research data is well-described and read for reuse by a third party. It 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.

Purdue Data Reuse Checklist

Data context

- Are the names and affiliations of all researchers and associated with the data?
 - Is contact information listed?
- Have you described when, how, and the tools used for data collection?
- Have you included versioning information such as release notes or improvements/change logs from the last version

Data structure

- Is your data's structure adequately described?
- Is your file naming structure intuitive? If not, have you included additional information to ensure it is understandable?
 - Have you used formatted dates and times consistently?
- Have you included a data dictionary or codebook to help others understand your data structure?
 - If using tabular data, are your column headings meaningful and complete?
- Have you used a standard method to represent missing data?

Data reuse

- Have you verified the data's quality and accuracy?
- Do the data contain sensitive or proprietary information?
 - If yes, have you taken measures to protect the identity of research participants or taken steps to embargo the data until it is appropriate to release?
- Do you own the data?
 - If yes, have you indicated such and referenced an appropriate Creative Commons license? Do you need (or does your institution claim) a patent?
- Have you thought about how you will provide access to your data?