Purdue Libraries: Resources and Engagement in Data Science

Michael Witt
Purdue University, mwitt@purdue.edu

Follow this and additional works at: https://docs.lib.purdue.edu/lib_fspres

Part of the Library and Information Science Commons

Recommended Citation
https://docs.lib.purdue.edu/lib_fspres/136

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 Libraries: Resources and Engagement in Data Science

Integrative Data Science Initiative Coffee Klatch - March 20, 2019

Michael Witt, Associate Professor and Head, Distributed Data Curation Center (D2C2)
<table>
<thead>
<tr>
<th>Library Subject</th>
<th>Librarian Name</th>
<th>Phone</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural Economics</td>
<td>Howard, Heather A</td>
<td>49-48164</td>
<td><a href="mailto:howar198@purdue.edu">howar198@purdue.edu</a></td>
</tr>
<tr>
<td>Agricultural Sciences Education &amp; Communication</td>
<td>Walker, Danielle N</td>
<td>49-67680</td>
<td><a href="mailto:walker524@purdue.edu">walker524@purdue.edu</a></td>
</tr>
<tr>
<td>Agriculture &amp; Biological Engineering</td>
<td>Walker, Danielle N</td>
<td>49-67680</td>
<td><a href="mailto:walker524@purdue.edu">walker524@purdue.edu</a></td>
</tr>
<tr>
<td>Agronomy (Soil science)</td>
<td>Walker, Danielle N</td>
<td>49-67680</td>
<td><a href="mailto:walker524@purdue.edu">walker524@purdue.edu</a></td>
</tr>
<tr>
<td>American History</td>
<td>Chapman, Bert</td>
<td>49-42837</td>
<td><a href="mailto:chapmanb@purdue.edu">chapmanb@purdue.edu</a></td>
</tr>
<tr>
<td>Animal Science</td>
<td>Walker, Danielle N</td>
<td>49-67680</td>
<td><a href="mailto:walker524@purdue.edu">walker524@purdue.edu</a></td>
</tr>
<tr>
<td>Animal Science</td>
<td>Roark, Kendall L</td>
<td>49-42637</td>
<td><a href="mailto:roark8@purdue.edu">roark8@purdue.edu</a></td>
</tr>
</tbody>
</table>
Taking a lifecycle view of data science:
Selected courses currently being offered

- BCHM 612 Bioinformatic Analysis of Genome Scale Data
- BCHM 495 R for Molecular Biosciences
- ILS 595 Data Management at the Bench
- ILS 295/ENGR 103 Developing Your Data Mind
- ILS 995 Data Sharing and Publication
- ILS 295/CS 290 Data Science and Society: Ethical, Legal, and Social Issues
- ILS 595/ANTH 592 Data Management and Curation for Qualitative Research
- HON 399 Critical Data Studies
- ILS 595-H Open Seminar in Critical Data Studies
- ILS 295 Data Management
- ILS 295 Introduction to Data Management *IDSI education ecosystem award
- ILS 695 Digital Humanities *IDSI education ecosystem award
- ILS 695 Digital and Analog Archives
- PHIL 293 Ethics for Data Science
Learning Communities

- Engineering in the World of Data
- The Data Mine: Critical Data Studies
Selected seminars and events

- Women in Data Science
- Critical Data Studies Lecture Series
- Dawn or Doom
- GIS Day
- Center for the Science of Information Summer Workshop
- The Tinkering Humanist
- Graduate student workshops

WOMEN IN DATA SCIENCE
GREATER LAFAYETTE @PURDUE

GIS Day

DAWN OR DOOM 2018
EMERGING TECHNOLOGY – RISKS AND REWARDS
# Data Storage at Purdue: Purdue Supported Storage

This guide details the variety of data storage options available to faculty and graduate students at Purdue University.

## Purdue Supported Storage

### Full guide to storage options at Purdue in printable form

- [Printable Guide to Storage Options](#)

Print this for a full listing of all storage options and facets

## Brief Introduction/Download

<table>
<thead>
<tr>
<th>Box</th>
<th>PURR</th>
<th>OneDrive/Office 365</th>
<th>Data Depot</th>
<th>Fortress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Free to Purdue affiliated researchers</td>
<td>10 GB free for 3 years; 100 GB for 10 years with grant; Extra project space is $1.08 per GB</td>
<td>Available to students with tuition</td>
<td>100 GB free trial option; Purchase in TB increments, currently $75 per TB per year</td>
</tr>
<tr>
<td>Available</td>
<td>1000 GB</td>
<td>1 TB</td>
<td>No upper limit; Total storage</td>
<td>No upper limit; Total storage</td>
</tr>
</tbody>
</table>
Sensitive Research Data Management: Guides, Articles and Other Resources

A guide for addressing issues with sharing research data involving human subjects or other sensitive data sets.

Guides

- **Secure Purdue - Data Classification & Handling**
  provides information on security policies, standards, data handling guidelines and ITSP security services

- **HIPAA Privacy**
  HIPAA Privacy homepage managed by Purdue University Student Health (PUSH)

- **Export Controls and Research Information Assurance (EVPRP)**
  Maintained by the Export Control and Information Assurance team within the Office of Research and Partnerships.

- **Confidentiality - CESSDA Council of European Social Science Data Archives**
  A guide for researchers seeking to collect and use sensitive data.

- **Guide to Social Science Data Preparation and Archiving (ICPSR)**
  Guidance from The Interuniversity Consortium for Political and Social Research (ICPSR), which is a repository of social science data, on data management best practices to prepare data for eventual sharing and archiving.
  - Chapter 5. **Final Project Phase: Preparing Data for Sharing** addresses confidentiality issues specifically.
  - Chapter 3. **Data Collection and File Creation Phase** addresses issues of coding schemes and other best practices for data handling.
  - Other useful documents from ICPSR:
Purdue University Research Repository

1. Helps you write an effective **data management plan** to include with your grant proposal
2. Gives you a private, online **project** space to share and manage your data with your collaborators
3. Enables you to **publish** your data in a scholarly context with a Digital Object Identifier (DOI)
4. **Archives** your data for a minimum of 10 years with standards-based policies and archival best practices
5. Reports monthly **impact** and how many times your data are viewed, downloaded, and cited
Research Data Management for Purdue

The Purdue University Research Repository (PURR) provides an online, collaborative working space and data-sharing platform to support Purdue researchers and their collaborators.

DOI: 10.4231/R7WH2N0H
Acquiring and managing collections of data...

To request a dataset to be purchased by the Libraries, speak with your subject librarian.

Some data collections and databases that are subscribed, purchased, or managed by Libraries include:

- GIS Portal
- Purdue e-Archives
- PURR
- ...and ~600 more from the Databases link on the Libraries main web page!
Links

- Purdue Libraries and School of Information Studies
- Librarians by subject
- Research Data division of Libraries (STEW 174)
- Spring 2019 ILS Course Offerings
- Engineering in the World of Data Learning Community
- The Data Mine: Critical Data Studies
- Women in Data Science
- GIS Day
- The Tinkering Humanist
- Library Guide: Data Storage at Purdue
- Library Guide: Sensitive Research Data Management
- Purdue University Research Repository
- Purdue GIS Portal
- Purdue e-Archives
- Libraries subscribed/purchased datasets and databases