

8-3-2020

## Research Data Service: Current State and Future Plan for CWU

Ping Fu

Central Washington University, ping.fu@cwu.edu

Follow this and additional works at: <https://digitalcommons.cwu.edu/libraryfac>



Part of the [Collection Development and Management Commons](#), and the [Scholarly Publishing Commons](#)

---

### Recommended Citation

Fu, Ping, "Research Data Service: Current State and Future Plan for CWU" (2020). *Library Scholarship*. 82. <https://digitalcommons.cwu.edu/libraryfac/82>

This Article is brought to you for free and open access by the James E. Brooks Library at ScholarWorks@CWU. It has been accepted for inclusion in Library Scholarship by an authorized administrator of ScholarWorks@CWU. For more information, please contact [scholarworks@cwu.edu](mailto:scholarworks@cwu.edu).

# **Research Data Service: Current State and Future Plan for CWU**

Ping Fu

Professor of Library Services  
Science & Data Management Librarian  
Head of Library Technology Services  
8/3/2020

## **Table of Contents**

Executive Summary

1. Introduction
2. Research Methods
3. Example Peer Institutions' RDS
4. Findings and Observations
5. Current Resources and Needs
6. What RDS Would CWU Libraries Provide?
7. Costs and Budgets
8. Recommendations
9. Conclusions
10. References
11. Appendix I – RDS survey questionnaire
12. Appendix II- Data Services Librarian's Responsibilities & Qualifications
13. Acknowledgements

## **Executive Summary**

Today, more and more academic libraries provide or plan to provide library-based research data services (RDS). There are two major reasons: (1) science becomes more collaborative, data-intensive, and computational, and academic researchers face a series of data management needs; (2) research funding agencies require researchers to provide data management plans when they apply for a grant and publishers also require researchers to provide data when publishing research results.

Tenopir, C. et al. (2014) pointed out that “academic libraries may be ideal centers for research data service activities on campuses, providing unique opportunities for

academic libraries to become even more active participants in the knowledge creation cycle in their institution.”

This study randomly selected some peer institutions and examined their “Research Data Services (RDS)” models to provide a baseline assessment of the current state of and future plans for research data services in the James Brooks Library at Central Washington University (CWU).

Example peer institutions selected for this report are: University of Washington Libraries (UW Libraries), Stanford University Libraries, and Western Washington University Libraries (WWU Libraries). UW Libraries and WWU libraries were selected because UW and WWU are peer institutions in Washington State and are members of Orbis Cascade Alliance (OCA). UW is a flagship university in OCA and Washington State while WWU is a regional comprehensive university comparable to CWU. Stanford University Libraries was chosen because Stanford libraries’ RDS covers all areas of RDS.

### **Key findings and observations:**

- A typical RDS model consists of four major parts: (1) a unit or a working group that provide RDS service; (2) a web portal or a platform to make resources and tools available, these resources and tools include their institutional repository, storage options, data management tools, data visualization tools, etc. (3) research guides, educational classes and drop-in workshops; and (4) providing metadata service and consultation service

### **What resources and tools we have:**

- Currently, we have an institutional repository, ScholarWorks, for faculty and students to deposit their scholarly works and publish preprint, post print, and grey print; some subject librarians have developed some research guides separately on their subject areas; very limited tools available.

### **What RDS Would CWU Libraries Provide?**

CWU Libraries would plan to support our researchers throughout the lifecycle of their research including providing support in finding resources to inform and enhance their projects; managing the details of their research and publishing processes; creating graphics to clearly communicate their findings; fulfilling funder requirements for data and research outputs; and elevating the visibility and impact of their work.

## **Recommendations:**

Based on the outline created by the Dean of Libraries, three recommendations are made:

- Recommendation for year one: working with the resources we have. Develop basic RDS model. Select one or two academic departments such as Geological Sciences and Biological Sciences as starting point.
- Recommendation for year two: If we hypothetically have a dedicated faculty position, extend RDS to all departments in the College of the Sciences. Consider collaborating on applying for a grant to enhance the RDS model.
- Recommendations for year three: scaling up and grant work. Continue to develop a mature RDS model and extend RDS to all colleges on campus.

## **1. Introduction**

A university library's Research Data Services (RDS) supports the research data management needs of faculty and students at the university by helping finding and accessing data, data management planning, data organization, reuse of data, data sharing and storage, data citation, and more. While many research libraries focus on providing a central repository for RD and involving in all cycles of research data management, small and mid-size libraries focus on promoting their institutional repositories, making tools available, developing research guides, and providing educational classes, drop-in workshops, and consulting services. Although RDS is a trend for academic libraries, Cox A.M., et al. (2019) found that most RDS development can be explained as the extension of traditional library services to research data.

## **2. Research Methods**

This study randomly selected some peer institutions and examined their "Research Data Services (RDS)" models to provide a baseline assessment of the current state of and future plans for research data services in the James E. Brooks Library at Central Washington University (CWU).

## **3. Example peer institutions' RDS**

### **3.1 University of Washington Libraries**

(<https://www.lib.washington.edu/dataservices>)

UW Libraries RDS supports the research data management needs of faculty and students at the University of Washington. This includes help finding and accessing data, data management planning, data organization, reuse of data, data sharing and storage, data citation, and more.

### **3.2 Western Washington University Libraries**

(<https://library.wvu.edu/use/scholarship>)

WWU libraries Scholarship & Research Services focuses on copyright issues, open access, institutional repository and research materials. It is still at an early stage of RDS.

### **3.3 Stanford University Libraries** (<https://library.stanford.edu/research/data-management-services>)

Stanford Libraries RDS covers all areas of RDS to help Stanford researchers with the organization, management, and curation of research data to enhance its preservation and access now and into the future.

## **4. Findings and Observations**

A typical RDS model consists of four parts:

**4.1 Libraries usually create a department/unit** consisting of a number of librarians and specialists to provide and support data services.

**4.2 Libraries usually create a web portal or a platform** to make resources, tools and services available. The resources and tools include their institutional repository, storage options, data management tools, data visualization tools, licensed data or free data etc. These tools assist researchers with data management planning and implementation, through training, consultations, resources, and guidance and help researchers throughout the research lifecycle. Here list some typical tools and data:

- (A) DMPTool is recommended for creating a Data Management Plan. <https://dmptool.org/> (from CA Digital Library)
- (B) Open Science Framework is recommended for managing research project and as a storage option.
- (C) Open Researcher and Contributor Identification (ORCID) for researchers.
- (D) Libraries' institutional repositories can be used for researchers to archive and share their data, preprint, etc.
- (E) Data visualization tool such as Tabeleau.  
<https://www.tableau.com/>
- (F) Libraries help create persistent identifiers or persistent archived versions of web pages. Libraries have institutional accounts for both of these services: [perma.cc](https://perma.cc) and [DataCite](https://datacite.org/) (DOI).
- (G) Research Electronic Data Capture ([REDCap](https://www.researchprotocols.org/2019/1/e14721)).
- (H) Licensed GIS data, GIS Software & Training, map collection, other free data, government data, etc.
- (I) Research labs or research commons.

**4.3 Libraries usually provide research guides, educational classes and workshops.**

(A). Data Management Guide including libraries' introduction to Data Management. Researchers can find information relevant to all stages of the research lifecycle - from Data Management Plans in the grant-writing process, to Organization and Format during data collection, to Sharing and Archiving their data post-publication. Libraries also provide a FAQ page for quick reference.

(B). A working group meets regularly and hosts seminars to discuss the issues and practices around tools and practices to enhance data sharing, preservation, provenance, and reproducibility.

(C). Open Access Guide for Open Access policy and open access in general.

(D). Provide in-class or drop-in workshop & consultation sessions to teach and help researchers better manage research data and preserve it for long-term use. Topics include data management planning, funder requirements for data sharing, metadata, tips to help keep research organized, sharing/archiving/preservation, and an introduction to tools and on-campus support.

#### **4.4 Libraries usually provide consulting services and metadata service**

(A). Data management plans and funder requirements including reviewing DMP or getting started with DMPTool to create a new plan; modifying this template as needed if researcher's funder requires a verification letter to use institutional repository as a sharing and preservation platform; identifying funder's requirements for data sharing and data management.

(B). Data management consultation including consulting on strategies and workflows for all stages of the research lifecycle, including storage and backup, data organization and description, workflow management software, sharing and archiving.

(C). Workshops and lectures on any aspects of the research data lifecycle.

(D). Data reference services including finding or acquiring data, identifying other campus departments that can help researchers analyze their data, and working with researchers to prepare their data for visualization.

(E). Persistent Identifiers including registering researchers' data for digital object identifiers (DOI) via DataCite to help researchers reference their own work in publications and allow others to cite their objects, and creating persistent snapshots of webpages at a moment in time via PermaCC.

(F). Metadata Creation helps researchers document and describe their data through metadata and other documentation techniques.

## **5. Resources and Needs We Have at CWU**

### **5.1 What resources do we have?**

- (A) ScholarWorks
- (B) Data resources including science data, maps & geospatial data, datasets, government data, open data, etc.
- (C) A number of research guides on data resources
- (D) Scholarly Communications Librarian & Subject Librarians
- (E) Other campus resources: Pivot, Excel, Adobe, Powerpoint.

### **5.2 What departments need data services?**

Among departments at the College of the Sciences, Geological Sciences, Biological Sciences, Computer Science, and Mathematics could be our targeted departments that need data services at the first beginning; however, we can collaborate with Graduate School and Grant Office to conduct a survey to have a clearer picture and help us pick one or two from the start.

## **6. What RDS Would CWU Libraries Provide?**

CWU Libraries would plan to support our researchers throughout the lifecycle of their research:

- (A) Find resources to inform and enhance their projects: Providing support for finding resources such as searching Pivot for finding research funding, datasets research guides created and/or maintained by Data Services Librarian and other subject librarians, and research guides for text mining resources.
- (B) Manage the details of their research and publishing processes: providing support for Data Management Plans and file organization strategies.
- (C) Create graphics to clearly communicate their findings: Providing support on GIS & geospatial services, data visualization, poster design and presentation design.



(D) Fulfill funder requirements for data and research outputs: Providing support in data sharing, preservation, citations, bibliographies, storages, etc.

(E) Elevate the visibility and impact of their work: Providing support in publishing researchers work in ScholarWorks, DOI, etc.

## **7. Costs and Budgets**

Cost estimates include salary and benefit for a new tenure track Data Services Librarian at Assistant Professor level, storages for data preservation, and software such as Research Data Management tools (e.g. DMPTool, Jupyter Notebooks), GIS & Geospatial Services tools (e.g. ArcGIS/ESRI Suite, Maptitude, PolicyMap, QGIS, Simply Analytics, Social Explorer, Story Map Json, OpenStreetMap), Data Visualization tools (e.g. D3.js, Excel, R/Rstudio, Tableau) and presentation and graphics tools (e.g. Adobe Illustrator, Adobe InDesign, PowerPoint). Strongly suggest university IT or Provost Office fund for purchasing/licensing these tools. Other costs include furniture, equipment, skills training, etc.

## **8. Recommendations**

The recommendations will be made based on the outline provided by the Dean of Libraries.

### **8.1 Recommendations for year one**, working with the resources we have.

For year one, consider conducting a survey to better understand needs; creating a working group, creating a RDS model prototype; creating a web portal to bring ScholarWorks, data resources, research guides, and data services together so we provide a centralized information platform for researchers and students; creating more research guides on Research Data Management, Data Management Plan and funder requirements and develop some classes and drop-in workshops on Data Services; creating a library faculty position on Science and Data Services; creating a consultation model to provide basic data services to one or two academic departments: Geological Sciences and Biological Sciences; building a partner relationships with Graduate Studies, university IS, and academic departments.

### **8.2 Recommendations for year two**, if we hypothetically have a dedicated faculty position.

For year two, consider maturing the RDS model based on the prototype we create at year one; recruiting a Science/Data Services Librarian; extending RDS services to the College of the Sciences; creating a RDS

unit. Consider collaborating with campus units such as university IT, academic colleges/departments to apply for a grant in enhancing the RDS model.

### **8.3 Recommendations for year three, scaling up, grant work.**

For year three, consider scaling up our RDS to all researchers and departments on campus to provide a full range of service on RDS; purchasing or leasing larger storage options if there are needs; purchasing more tools, either storage or visualization or data management tool if necessary; deeply involving in grant work and/or data management plan.

## **9. Conclusions**

This white paper simply examined some existing RDS models, summarized major components of a typical RDS model, illustrated what RDS that CWU Libraries would provide, and estimated cost and budget for the suggested RDS model. We also analyzed our current state and then made three phases recommendations based on the outline illustrated by the Dean of Libraries.

## **10. References**

1. Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(2), 84-90.
2. Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. (2019). Maturing research data services and the transformation of academic libraries. *Journal of Documentation*, 75(6), 1432-1462.
3. Joo, Soohyung, & Peters, Christie. (2020). User needs assessment for research data services in a research university. *Journal of Librarianship and Information Science*, 52(3), 633-646.

## **11. Appendix I – RDS Survey Questionnaire**

The survey is designed to help understand the following questions: what types of data-related assistance have CWU researchers received from the university? What types of data-related resources do researchers need for their research activities? To what extent are they willing to use these resources? How do researchers currently manage their data, e.g. data

documentation and sharing? We adopted the survey questions from Joo, Soohyung, and Christie Peter's research.

1. What is your research field?
  - a. Agriculture, Food and Environment
  - b. Humanities & Arts
  - c. Social Sciences
  - d. Sciences & Engineering
  - e. Medicine/Health Sciences
  - f. Business
  - g. Education
  - h. Others
2. What is your status?
  - a. Assistant professor
  - b. Associate professor
  - c. Full professor
  - d. Full time researcher/post-doc fellow
  - e. Senior lecturer
  - f. Lecturer
  - g. Graduate student
3. What is your age group?
  - a. Under 24
  - b. 25-34
  - c. 35-44
  - d. 45-54
  - e. 55-64
  - f. 65 or older
4. Have you received assistance on campus with *type of assistance*?
  - a. Data management planning
  - b. Data collection
  - c. Data refinement/cleaning
  - d. Data analysis
  - e. Data documentation
  - f. Data visualization
  - g. Finding existing datasets for data collection
  - h. Finding an appropriate repository for long-term storage
  - i. Meeting funder mandates

5. If yes, assistance from people in which area?
  - a. Department
  - b. College
  - c. Library
  - d. Campus IT
  - e. Research Center
  - f. Other
6. Rate the level of support that you need for the following research processes and activities
  - a. Assistance with data management plans
  - b. Assistance with data collection
  - c. Assistance finding existing datasets
  - d. Assistance finding an appropriate repository for data archives
  - e. Assistance with data refinement or cleaning
  - f. Assistance with “qualitative” data analysis
  - g. Assistance with “quantitative” data analysis
  - h. Assistance with data documentation
  - i. Assistance with data visualization
  - j. Assistance with meeting funder mandates for sharing research data
7. Which format of data do you generate in your research?
  - a. Tabular or spreadsheet data (e.g. CSV, MSExcel)
  - b. Textual data (e.g. TXT, MSWord)
  - c. Image files (e.g. JPEG, TIFF)
  - d. Audio files (e.g. WAV, MP3)
  - e. Video files (e.g. AVI, WMV, MP4)
  - f. Geospatial data
  - g. Artifacts, samples, and/or specimens
  - h. Semi-structured format data (e.g. XML, JSON)
  - i. Other
  - j. Genomic data
8. Where do you store your data?
  - a. Computer or external storage (external hard-drive, flash drive, CD/DVD)
  - b. Lab computer or storage
  - c. Department or college server

- d. University server
  - e. Library repository
  - f. Cloud storage (e.g. Dropbox, Google Drive)
  - g. External data repository
  - h. Other
9. Which of these options approximately describe your current digital storage needs for research data and related materials?
- a. Less than 1GB
  - b. 1-10GB
  - c. 10-100GB
  - d. 100GB-1TB
  - e. 1TB-10TB
  - f. 10TB-50TB
  - g. 50TB+
10. Do you intend to share the data sets associated with your research?
- a. Willing
  - b. Not willing
  - c. No response

**12. Appendix II - Data Services Librarian's Responsibilities & Qualifications**

The following description of Data Services Librarian's responsibilities \* qualifications was adopted from the Data Services Librarian job advertisement of Clemson University Libraries on internet.

**A. Responsibilities:**

**Data & Research Services**

- Provides support and instruction in data management tools and techniques required by faculty, researchers, and students to find, describe, preserve, and visualize data.
- Assists researchers with the review of data management plans and development of grant proposals.
- Analyzes and assesses campus data management needs, current initiatives, and future directions.
- Develops and implements plan for the Libraries to provide and sustain data services for the University community.

- Works collaboratively across the Libraries to support various initiatives including Open Access, Open Educational Resources, and digital initiatives.
- Develops partnerships and works in collaboration with campus offices that provide research data services to support the University's research mission.
- Engages in professional development such as attending conferences, workshops, and webinars related to job functions.
- Develops a focused program of high quality research and creative accomplishments, consistent with professional responsibilities and the Libraries' mission and goals.

### **Instruction & Outreach**

- Designs and teaches data management instruction sessions, workshops, and programs; creates instructional materials and research guides in a variety of formats.
- Participates in outreach to promote library resources and services.

### **Service**

- Actively participates and demonstrates leadership in professional responsibilities that serve the Libraries, University, profession, and community

### **B. Qualifications:**

- An ALA-accredited graduate degree in librarianship or other relevant graduate degree as deemed appropriate by the Libraries
- Demonstrated knowledge of data management practices, data curation practices, and preservation principles and practices
- Understanding of the research life-cycle as demonstrated by education or work experience
- Effective interpersonal, oral, and written communication skills
- Demonstrated ability to work in a team environment
- Experience or demonstrated potential supporting researchers with data services, including review of data management plans, data discovery, and data storage
- Familiarity working with common funding agency requirements, such as NSF, USDA, NEH, NIH
- Familiarity with software used for statistical and/or qualitative analysis and data visualization
- Demonstrated understanding of the value of diversity and inclusion in the workplace

### **13. Acknowledgements**

This work would not have been possible without the support of the CWU Library Administration and all Library Faculty members. I am especially indebted to Ms. Rebecca Lubas, Dean of Libraries, who provided general guideline, outline and positive feedback for this white paper, Ms. Elizabeth Brown, Associate Professor and the Chair of Library Faculty, who provided constructive suggestions and feedback to this white paper, and Aimee Quinn, Assistant Professor and Social Science Librarian, who spent her time in proofreading and editing for this whitepaper.