Apr 9th, 10:45 AM

DataShare: Lessons from the Field

Julia Kochi  
*University of California, San Francisco*

Megan Laurance  
*University of California, San Francisco*

Follow this and additional works at: [https://escholarship.umassmed.edu/escience_symposium](https://escholarship.umassmed.edu/escience_symposium)

Part of the [Scholarly Communication Commons](https://escholarship.umassmed.edu/escience_symposium)

This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

[https://escholarship.umassmed.edu/escience_symposium/2014/program/2](https://escholarship.umassmed.edu/escience_symposium/2014/program/2)

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts and New England Area Librarian e-Science Symposium by an authorized administrator of eScholarship@UMMS. For more information, please contact [Lisa.Palmer@umassmed.edu](mailto:Lisa.Palmer@umassmed.edu).
DataShare: Lessons from the Field

Julia Kochi and Megan Laurance
University of California, San Francisco Library

E-Science Symposium
April 9, 2014
Today I’ll Cover...

- Overview of DataShare
- Publicity and Marketing
- Case Studies
- Lessons Learned & Next Steps
Today I’ll Cover:

- Overview of DataShare
- Publicity and Marketing
- Case Studies
- Lessons Learned & Next Steps
DataShare Overview

- Brainchild of Michael Weiner, M.D.
- Since then:

**FAIR ACCESS TO SCIENCE AND TECHNOLOGY RESEARCH ACT (FASTR)**

[Image of White House website]

[Image of PLOS logo]
Goal

- Simplify data deposit for UCSF researchers in order to facilitate widespread data sharing
  - User-centered
  - Easy to use interface
  - Self-service
  - Minimal metadata
  - Format and subject agnostic
Prompted to log in via campus account
Data Type

Use this field to specify the type of file(s) you will be uploading as part of this dataset. DataShare accepts all file formats. Generally speaking, they should fall into one of the following types:

- text
- spreadsheet (csv, tab-delimited, Excel)
- video
- image
- software code
- pdf

If your data files are of a specialized or proprietary nature, be sure to include any pertinent information that would allow the proper viewing and/or usage of the file(s), either as a separate 'Readme' document within the data set folder, or in the 'Technical Description'.

If the dataset consists of multiple files, we recommend, but do not require, that you put all of the files into a single compressed folder (e.g. zip, tar) before uploading it to DataShare.
Review before submitting to DataShare

Metadata
- Click here to view metadata

Record contains all required fields
Record contains all recommended fields

Go Back
and edit description

Uploaded Files
Record must include at least one file

Upload
more files

Ready to go?

- Dataset sent to DataShare repository
- Receive doi, citation
- Record available in MyDatasets
### My Datasets

<table>
<thead>
<tr>
<th>Title</th>
<th>Publisher</th>
<th>Publication Year</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-analysis of microarray data from rapamycin resistant DLBCL cell lines</td>
<td>University of California, San Francisco</td>
<td>2013</td>
<td>Edit</td>
</tr>
</tbody>
</table>

[Add new dataset](#)
Search for Data

Search bar for entering keywords and finding data.

Search bar:

Go

Or

» Browse all data

Share Data

Make your data count
Get credit for your data
Ensure reproducibility
Promote reuse
Meet funder requirements

» Learn more

Featured Dataset

Progression of white matter degeneration in amyotrophic lateral sclerosis: A diffusion tensor imaging study. White matter damage in frontotemporal dementia and Alzheimer’s disease measured by diffusion MRI. Patterns of age-related water diffusion changes in human brain by concordance and discordance analysis.

Researcher Voice

“Making data transparent and available is going to accelerate all of science. It’s a relatively inexpensive way to get more value out of all of the work that we do.”

-Dr. Michael Weiner, UCSF

DataShare was made possible through a partnership between the California Digital Library, UCSF Clinical & Translational Science Institute, and UCSF Library. © 2013 The Regents of the University of California
CSF Biomarker and PIB-PET Derived Beta-Amyloid Signature Predicts Metabolic, Grey Matter and Cognitive Changes in Non-Demented Subjects.

Today I’ll Cover...

- Overview of DataShare
- Publicity and Marketing
- Case Studies
- Lessons Learned & Next Steps
Publicity and Marketing

- Researchers with funder/publisher requirements
- Data Managers
- Lab Managers
- Recipients of OA funding

- OA Week events
- Group presentations
- Campus-wide tech fair
- Integrated into postdoc and graduate level courses
Today I’ll Cover...

- Overview of DataShare
- Publicity and Marketing
- Case Studies
- Lessons Learned & Next Steps
Case Study 1

“*We’re done with this data. We’ve picked all the fruit.*” - *Data Manager, Women’s Health Research Group*

- NIH study concluded, findings published
- Requests from other researchers
- No obvious home for data (e.g., subject repository)
I want to share my data, but my current method is costly. - Professor, Dept. of Biochemistry and Biophysics

- Repository currently using too cumbersome
- Actual cost vs time cost
- Dataset size
- Long-term storage
Case Study 3

❖ I need something for a grant I’m writing.

- Divisional Analyst, OBGYN & Associate Professor, Dept. of Medicine

• Immediate need to fulfill grant proposal
Other Use Cases

- Outside collaborators
- HIPAA compliance
- Continuously updated dataset
Today I’ll Cover...

- Overview of
- Publicity and Marketing
- Case Studies
- Lessons Learned & Next Steps
Lessons Learned

- It’s hard to get noticed
- It’s even harder to break existing patterns of data sharing
- User feedback is key
- Insert at point of pain/need
Next Steps

- Continue usability testing and user interviews
- More outreach and marketing
- Develop stronger partnerships
- Develop a deeper data management toolkit
- Participate in further development of DataShare at a UC-wide level
Acknowledgments

Mini Kahlon
Angela Rizk-Jackson
Anirvan Chatterjee
NCATS – NIH Grant # UL1 TR000004

Michael Weiner

Karen Butter
Julia Kochi
Megan Laurance
Geoffrey Boushey

Patricia Cruse
Carly Strasser
Perry Willett
Stephen Abrams
Marisa Strong
Eric Hetzner & Abhisheek Salve
Contact Info

datashare.ucsf.edu

Julia Kochi: julia.kochi@ucsf.edu

Megan Laurance: megan.laurance@ucsf.edu