



Data Sharing Made Easy:
The UC Research Data Publication Service

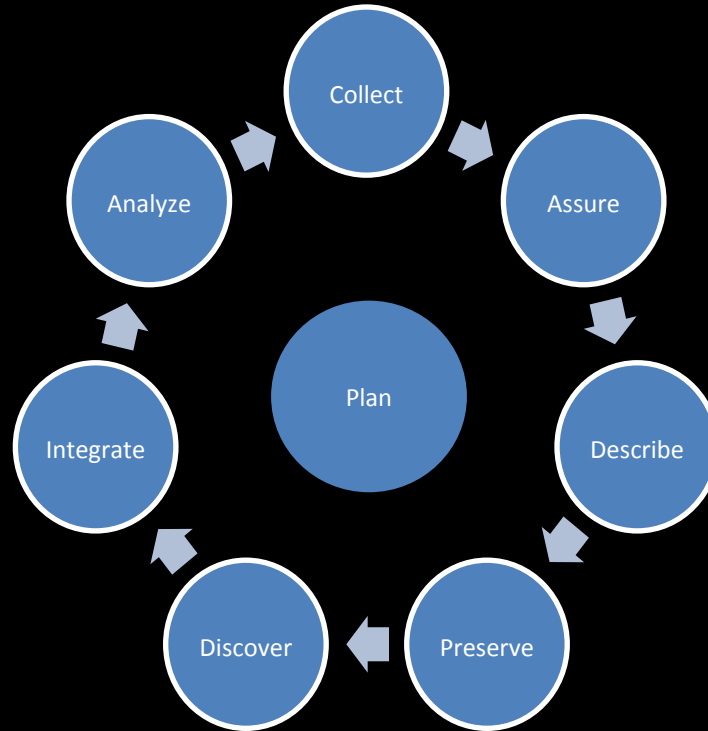
Marisa Strong
California Digital Library / UC Curation Center

California Digital Library

- ❖ 4th Floor of the Oakland Scientific Facility (OSF), Oakland, CA
- ❖ Office of the President - Academic Affairs
- ❖ National and global services for campuses, publishing, collection development, and curation-supported scholarship at UC



Scholarly lifecycle



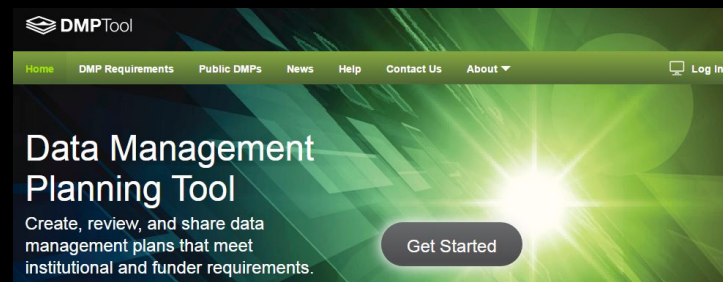
— Strasser et al. (2015), *A Primer on Data Management*
<escholarship.org/uc/item/7tf5q7n3>



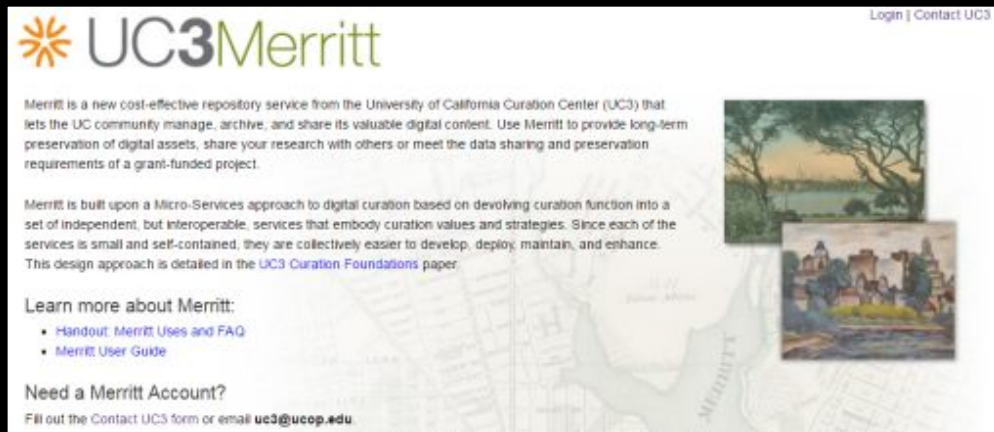
University of California Curation Center

UC3 provides services for scholarly preservation and reuse of curated digital content.

www.cdlib.org/uc3



dmptool.org



merritt.cdlib.org

Current version of dash.cdlib.org



Describe, manage, and share your research data

Explore Dash

[Find data](#)
[Learn more about Dash](#)
[Read FAQ](#)
[Visit Dash on GitHub](#)

Not part of the UC community?

[Go to DataONE Dash](#)

Visit UC Dash Sites

Find and share data from your UC Campus

[UC Berkeley](#)

[UC Irvine](#)

[UC Merced](#)

[UC Office of the President](#)

[UC Riverside](#)

[UCSF DataShare](#)


[UC Santa Cruz](#)

Don't see your UC campus? They are not currently participating. Contact us for more information.


Campus Collaboration

Geolocation support from UC Irvine Library

Orange County Data Portal
<https://dash.lib.uci.edu/xtf/search?smode=orangecounty-home>



UNIVERSITY OF CALIFORNIA



About
Search Data
Share Data
My Datasets
UCI Dash | OC Data Portal

Select a Dataset...

Publisher

- UC Irvine (3)
- UC Merced (2)

Author

- Allison, Steven D. (1)
- Bales, Roger (1)
- Bread, Scott (1)
- Dashwood, Marianne (1)
- Ebert, Tula (1)
- Gough, Susanne (1)

more

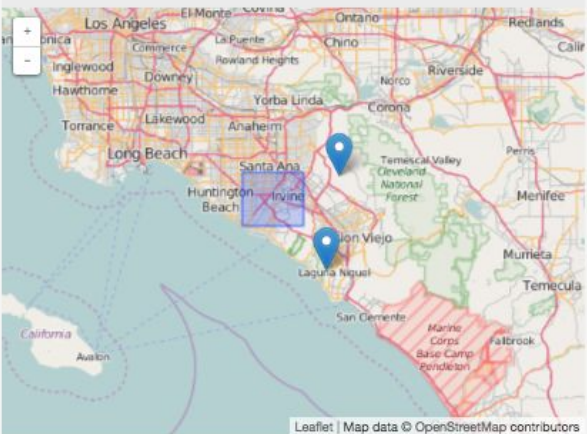
Keyword

- air temperature (2)
- snow depth (2)
- soil moisture (2)
- soil temperature (2)
- water balance (2)

more

[Clear search](#)

NOTE: Only data with assigned geoLocation values will appear in this map interface.



Leaflet | Map data © OpenStreetMap contributors

Can microbial functional traits predict the response and resilience of decomposition to global change?

by Allison, Steven D.
UC Irvine
> Show abstract

Awarded Alfred P. Sloan Grant in 2015

Problem: Researchers aren't archiving and sharing data in sustainable ways. Reasons being lack of self-service options, difficult to use, designed for publications, not research data

Solution: Allow for researchers to easily deposit data into community repositories owned and operated by trusted organizations. Provide ability to find, retrieve and reuse data.

Dash will be a well designed, user-friendly, data curation platform to be layered atop existing repositories.

Analysis of other frameworks available but they were too tightly coupled to other repositories

UC3 decided to design and implement our own utilizing standard protocols and open source software.



Agile Development Approach



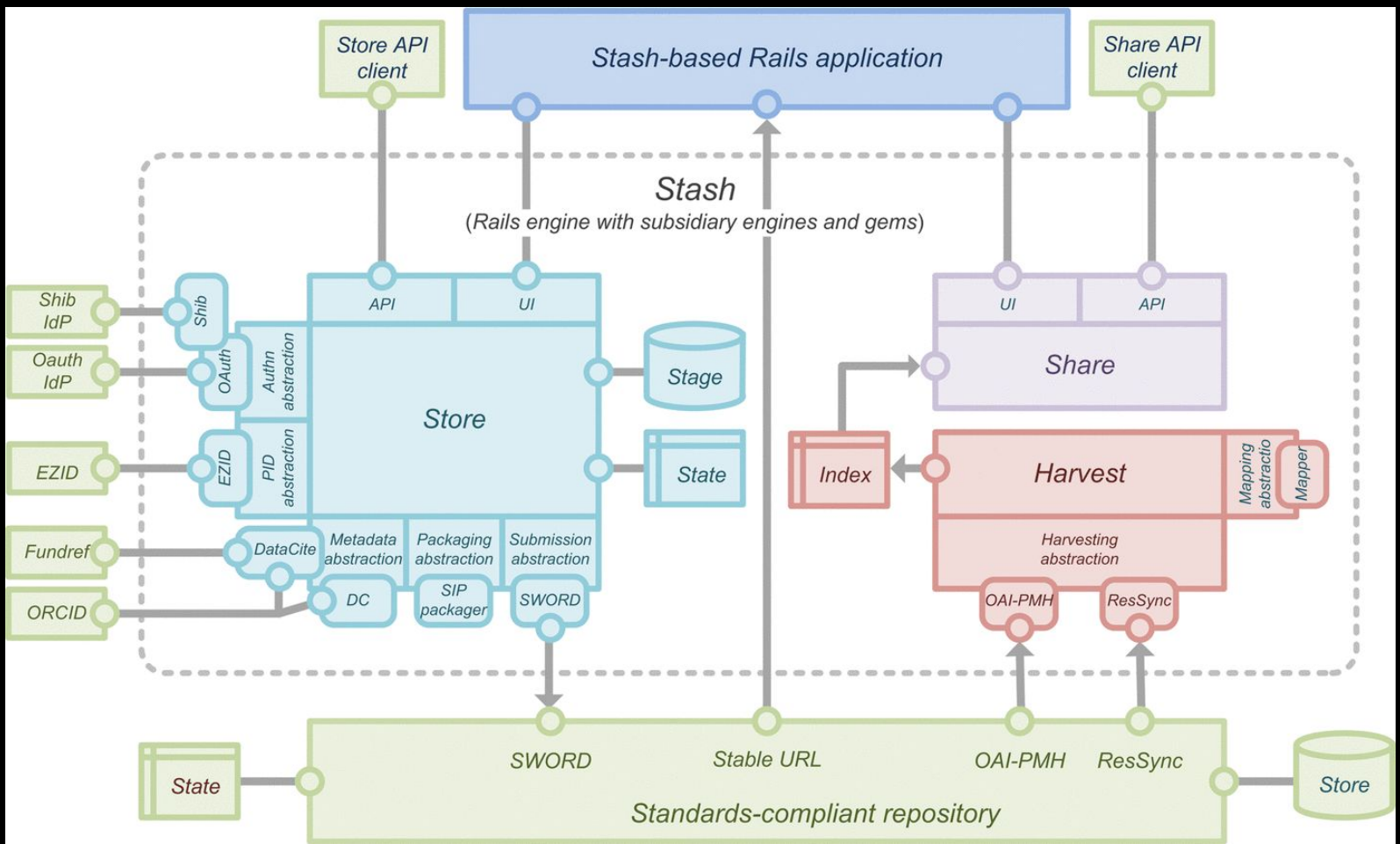
Dash provides self-service deposit of datasets

- Prepare
- Select
- Describe
- Identify
- Preserve
- Discover

Stash Framework provides architecture for the Dash application

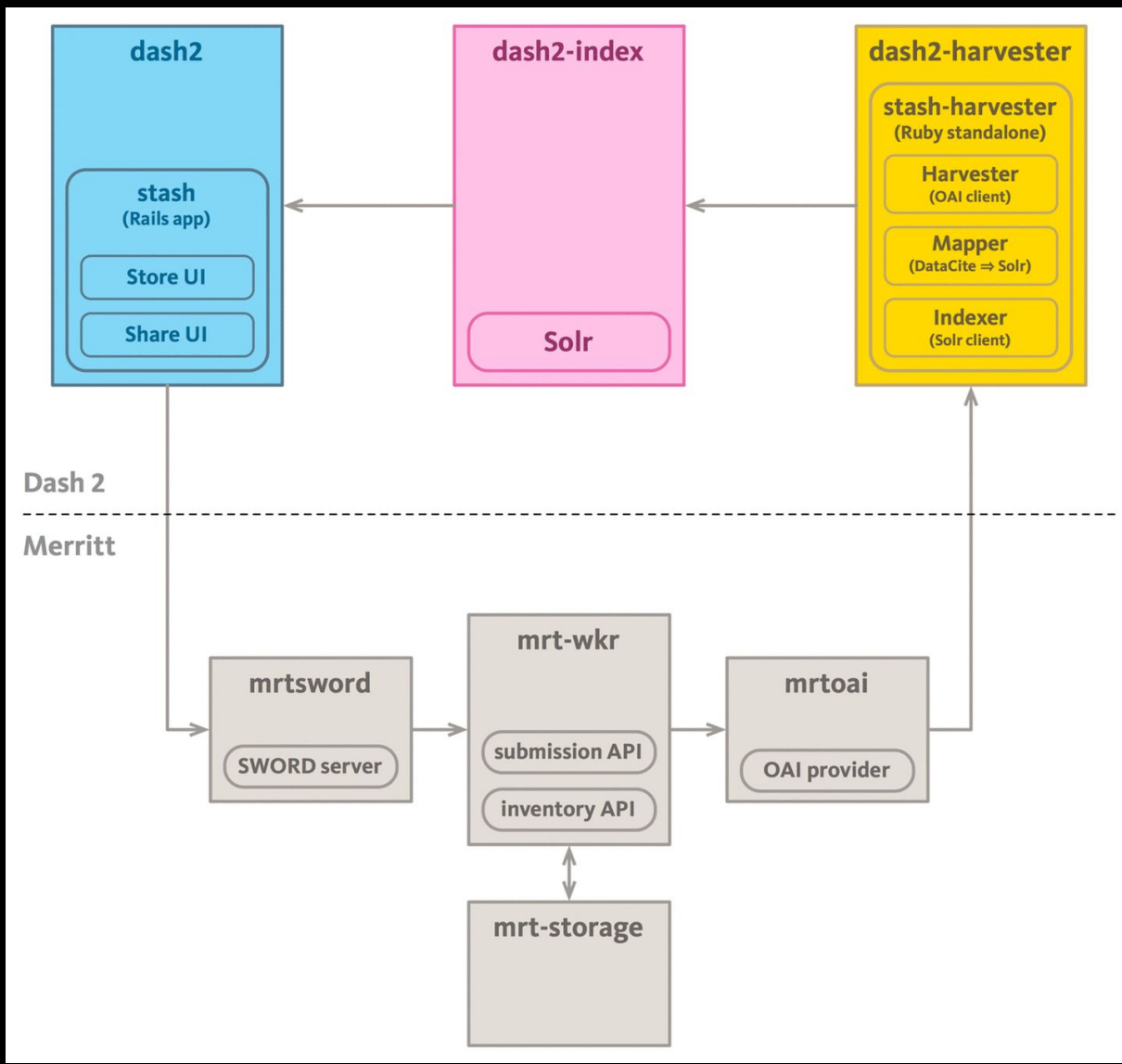
- pluggable architecture for additional protocols and metadata schemas
- works with your existing repository using standard protocols
- supports single- or multi-tenant deployment

Stash Framework



SWORD Protocol

- Dash curation platform is applicable to any repository supporting the SWORD protocol for submission and metadata harvesting.
- SWORD is itself an enhancement to the Atom publishing protocol (Atom-Pub, 2007).
- Merritt repository has been enhanced to conform to SWORD 2.0.
- <https://github.com/CDLUC3/stash-sword>



Rails Engines

Multiple engines and gems to plug into Stash framework

<https://github.com/CDLUC3/dashv2> - Rails application

https://github.com/CDLUC3/stash_engine

https://github.com/CDLUC3/stash_ezid_datacite

https://github.com/CDLUC3/stash_datacite

<https://github.com/CDLUC3/datacite-mapping>

<https://github.com/CDLUC3/dash2-harvester>

<https://github.com/CDLUC3/stash-harvester>

<https://github.com/CDLUC3/stash-wrapper>

Integration with UX & Web Production Team

Web Production - completely independent library of components and themes which can be incorporated into the Rails application - <https://github.com/cdlib/dash-ui>

Fully functional to test functionality

Utilizes Bower for package management and Gulp to handle repetitive tasks

Accessibility - incorporate UCOP best practices to support accessibility

[Dash UI Library](#)

Keywords Component

Keywords

apples x | here is a long keyword to investigate wrapping at small screens x | strawberries x | grapes x

Wireframes built with Axure

Logos

Q

[Home](#) | [Browse Data](#) | [Help](#) | [My Datasets](#) | [Logout](#) | [username]

[Title] submitted with doi:XXXXXXXXXX. There may be a delay for processing before the item is available. ✕

My Datasets

Start new dataset

In progress						
Title	Description	Files	Created	Modified		
Fast Charging Tests	required: 2 of 4 optional: 0 of 6	2	10/08/2014	09/23/2015	Resume	Delete
Maternity Leave Educational Tool Evaluation	required: 4 of 4 optional: 6 of 6	5	08/31/2014	12/04/2014	Resume	Delete
Annual Survey of Orange County 2000	required: 1 of 4 optional: 3 of 6	0	09/19/2014	05/26/2015	Resume	Delete
14c dates from Taraco, Peru	required: 4 of 4 optional: 1 of 6	1	10/22/2013	10/27/2014	Resume	Delete
Lensfree On-chip Microscopy for Telemedicine, v2	required: 3 of 4 optional: 4 of 6	0	12/06/2013	11/03/2014	Resume	Delete

[Show all](#)
◀ prev
1
2
3
next ▶

Multi-Tenant / Campus Configuration

- <https://github.com/cdlib/dash2-config>
- each campus has their own domain, provide by virtual hosts in Apache (dash.berkeley.edu).
- Campuses provide certificate information, CNAMEs that point to our servers, InCommon certificates, and Shibboleth metadata.
- Other configurable parameters:
 - repository collection profile
 - abbreviations & logos



Shibboleth Authentication for campuses (omniauth_shibboleth gem)

Google Authentication for DataONE (omniauth_google_oauth2 gem)

The screenshot shows the dash.berkeley.edu website interface. At the top left is the 'dash | BERKELEY Library UNIVERSITY OF CALIFORNIA' logo. The URL 'dash.berkeley.edu' is centered at the top. Navigation links include 'Home', 'Browse Data', 'Help', 'Login', and 'More Dash Sites'. A search bar is located at the top right. A 'Participating Partners' dropdown menu is open, listing: DataONE, University of California, Berkeley, University of California, Office of the President, and developer login (only development). A 'CalNet Authentication Service' overlay is centered on the page, featuring the Berkeley University of California logo and the following fields: 'CalNet ID:' with the value 'marisa.strong', and 'Passphrase (Case Sensitive):' with a masked input '.....'. Below these fields are 'SIGN IN' and 'HELP' buttons. At the bottom of the overlay, it says 'Copyright © 2016 UC Regents. All rights reserved.' The background of the website features a historical photograph of a field with the text 'data sharing made easy' and a quote 'but data remain' by Santiago Ramón y Cajal, 1897.

Describe Dataset

Upload Files

Review and Submit

Describe Your Datasets

Basic Information

Type of Data *

Spreadsheet

Title *

Author(s)


First Name

Last Name *

Institutional Affiliation *

x remove

+ Add Author

 Connect your ORCID ID



Abstract *

Describing your data

- Dash uses Datacite metadata schema due to it's minimal amount of required metadata, yet providing capabilities to provide rich metadata.
- Provides required data for DOIs (distinct object identifiers) offered by the EZID service.
- Assign geolocation data to describe where data was collected.
- ORCID integration - Oauth to authenticate and assign your own ORCID to you dataset. This in turn will be discoverable via ORCIDs and datacite?
- Fundref - lookup of funders provided by CrossRef

Data Description (optional) ▲

Funding

Granting Organization

Award Number

× remove

+ add another funder

Keywords

Methods How was this dataset collected? How has it been processed?

Usage Notes What else would someone need to know to use this dataset?

Related work(s) articles, other datasets: e.g., This dataset is cited by DOI 10.1000/182.

Relationship

ID Type

Identifier

× remove

+ add another related work

Describe Dataset

Upload Files

Review and Submit

Upload Your Files (optional)

Size limits: 2GB per file, 10 GB in total. To upload larger datasets, contact us.

Step 1: Choose your files

Drag and drop files here

or

Choose Files

Step 2: Upload your files

UCIDataPortalRequirements.xlsx 12.60 KB

Remove

Total: 12.60 KB

Upload

Back to Describe Dataset

Proceed to Review

Other

The Influence of Probabilistic Methodologies on Networking

Gil, Thomer M., Massachusetts Institute of Technology (MIT)

Citation

Gil, Thomer M. (2016), The Influence of Probabilistic Methodologies on Networking, v1, Other, <https://dx.doi.org/10.5072/FK2GQ6WZ4Z> (opens in a new window)

Abstract

In recent years, much research has been devoted to the exploration of von Neumann machines; however, few have deployed the study of simulated annealing. In fact, few security experts would disagree with the investigation of online algorithms. STEEVE, our new system for game-theoretic modalities, is the solution to all of these challenges.

Methods

Usage Notes

Keywords

Funding

References

 Download this dataset

Data Files

> Version 1

Metrics

 7
views

 0
downloads

Keywords

License

This work is licensed under a Creative Commons Attribution 4.0 International license.



Explore and discover...

Find the data you need

Find by...



Placename



Subject

ceperis, conurationem, fecerunt,
gracchum, magic, obscurein, roma, testing,
that, more »

Find by location



Stash Continuous Integration

Travis UI - <https://travis-ci.org/CDLUC3/dashv2>

- rspec unit tests and code analyzer - rubocop

Deployment - Capistrano - deployment of Rails engines

- Gem versions MUST BE defined properly in the Gemfile.lock file

Dev/Stage/Production servers

- Hosted on AWS
- Puppet for configuration management
- AWS RDS MySQL 5.6

Current state of Dash

Currently configured for UCOP, UC Berkeley, and DataONE

- dash2-stg.berkeley.edu
- dash2-stg.ucop.edu
- oneshare2-stg.cdlib.org

Usability testing for iterative improvements to service

Future of Dash

- Continuing Partnerships – UCSC, UCLA
- Embargo of datasets, Geolocation enhancements, UC Santa Cruz development support
- DATASHARE at University of Edinburgh / Digital Curation Centre
- Soon recruiting for a service manager to promote and support DASH within the UC scholarly community

Dash Project Team Members

Stephen Abrams: Architect

Marisa Strong: Technical Project Lead

Scott Fisher: Technical Lead/Ruby Developer

David Moles: Technical Lead/Ruby/Java Developer (Harvester, SWORD)

Bhavi Vedula: Ruby Developer (remote developer)

David Loy: Java Developer (SWORD)

UI/UX Team:

John Kratz: UX Design & **Joel Haggdorn:** Web Production

John Chodacki: Director of UC3

Perry Willett: Service Manager

University of California Curation Center (UC3)

Learn More

- dash.cdlib.org
- cdluc3.github.io/dash

