

# A decentralized scholarly commons

DANIELLE ROBINSON, PhD  
Co-Executive Director  
Code for Science & Society

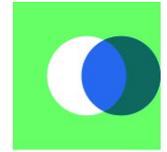
@daniellecrobins @codeforsociety

**CS&S**



# Code for Science & Society

Nonprofit, open source,  
in the public interest

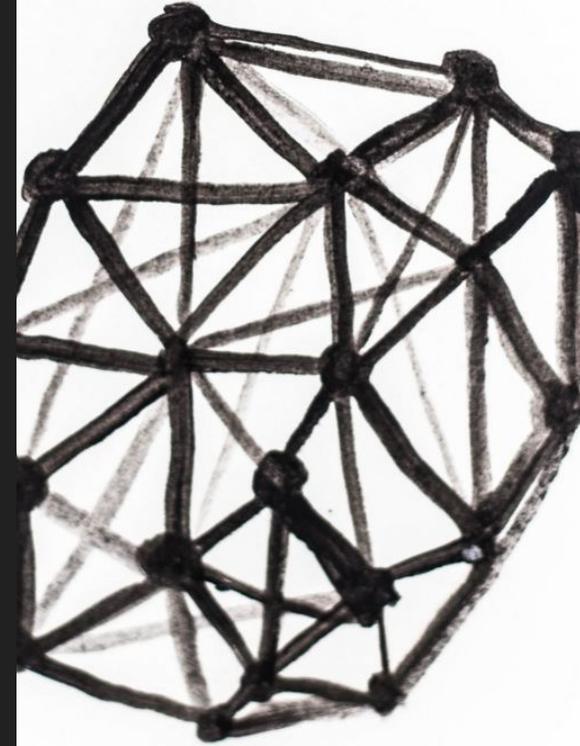


Stencila



CS&S

What would digital infrastructure that centers our values look like?



The web is today's  
scholarly commons

Early work of a writer

Government data

Newspaper archives

Your family photos

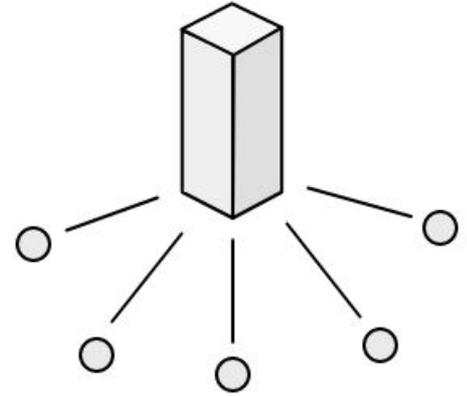
Scholarly data



Everything on the web  
is data

Era of megaplatforms

Our data = their business



Today's web lacks

Persistent identifiers

Transparent change log

Links between silos

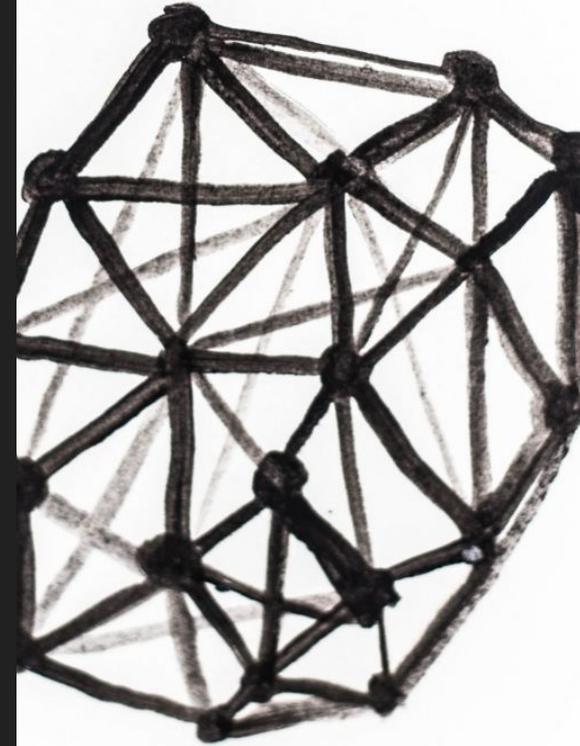




The web is being reimagined

**CS&S**

The time to center our  
values is now



Dat is

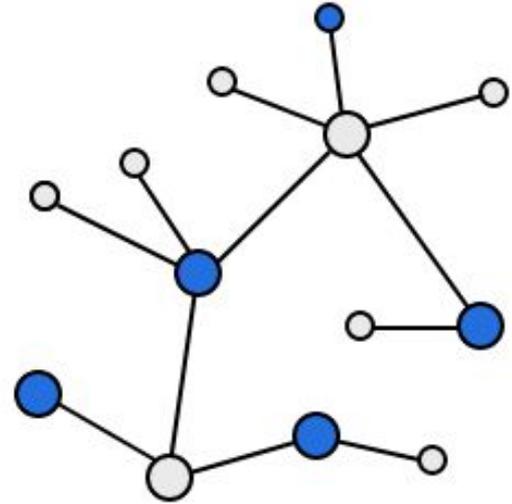
Persistent identifiers

+

Append-only log

+

Network of peers



**CS&S**

Yes this is starting to  
sound like

Utopian tech solutionism



Dat was developed for  
scientific data sharing

'Dat in the Lab'

Embed Dat developer  
inside a research group





# Start with data creation

Dr. Dannise Ruiz-Ramos describes sea star genome annotation pipeline



Researchers want

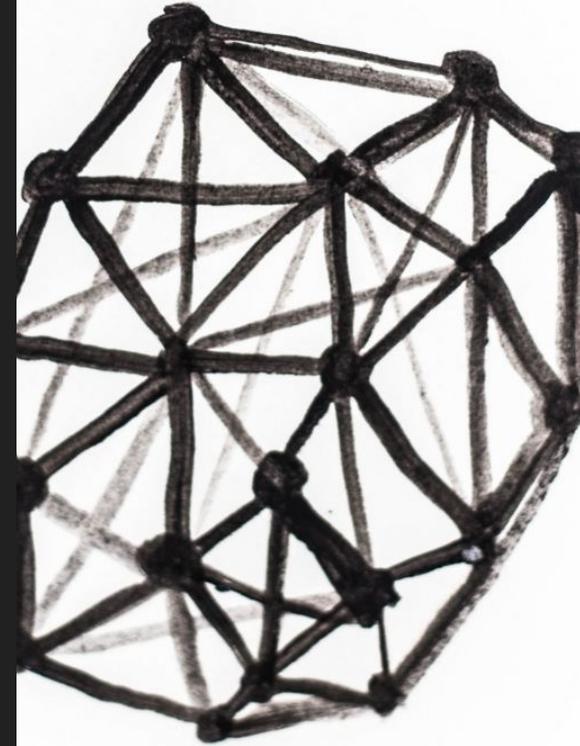
Data sharing, versioning

Data preservation + use  
metrics, demonstrate value  
of data

Low costs, high trust



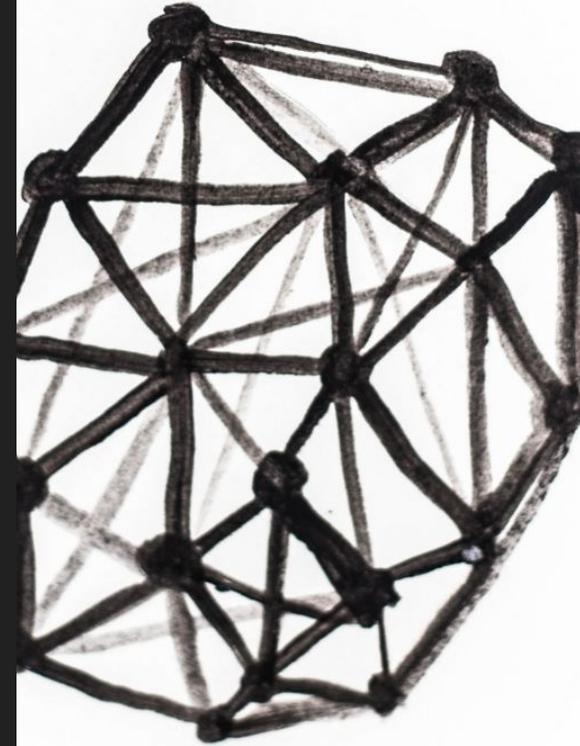
What does a  
decentralized scholarly  
commons look like?



A community-driven  
network

Built with partners

Designed for long term  
preservation and access



Collaborative approach

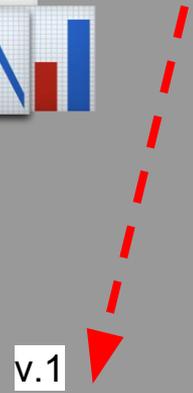
Link trusted institutions

Leverage existing  
infrastructure

Build a scholarly commons  
by linking silos







v.1





v.1



v.1



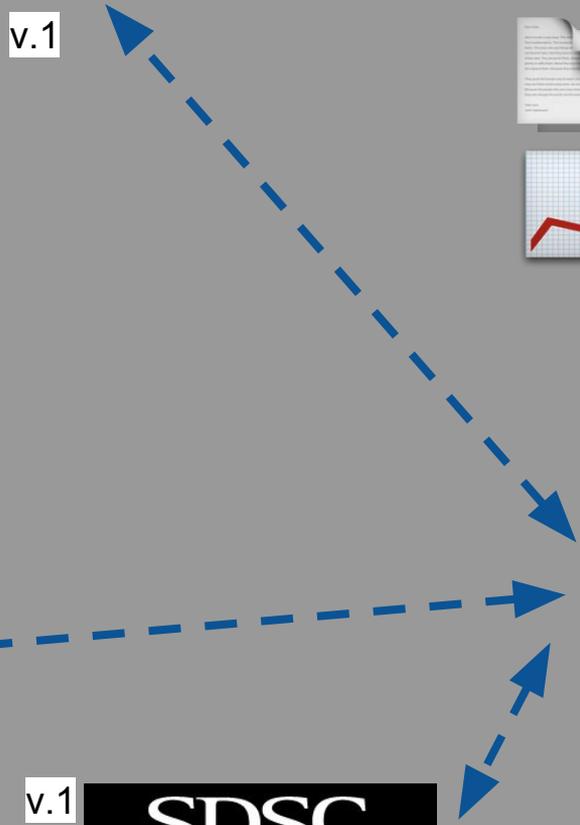
v.1

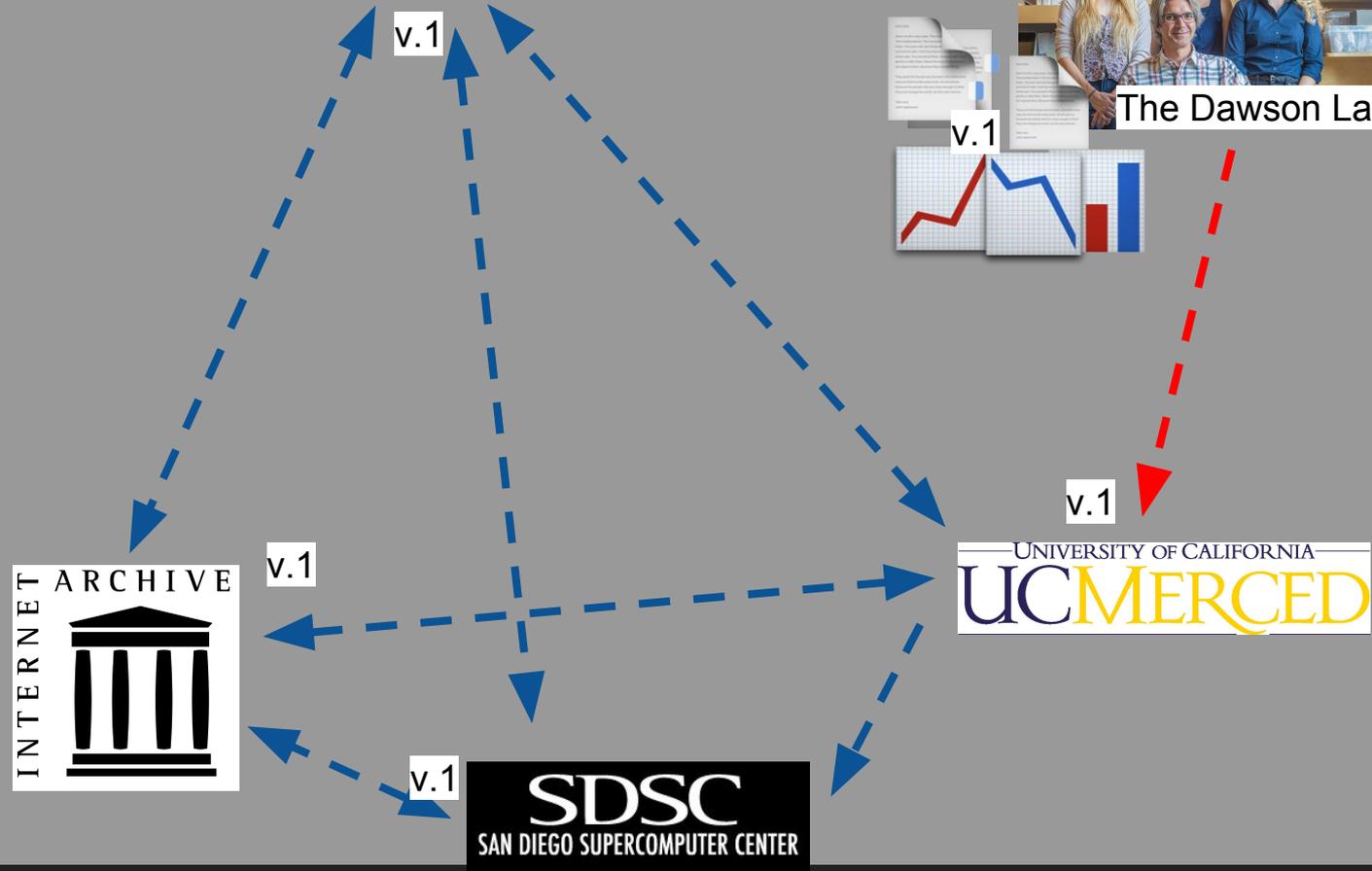


v.1



v.1







v.1

v.1

v.1

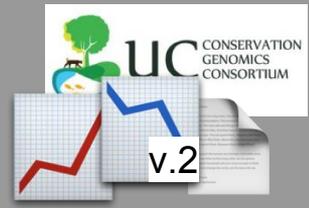


v.1



v.1





v.1

v.1

v.1

v.1

v.1

v.1





UC Curation Center



The Dawson Lab



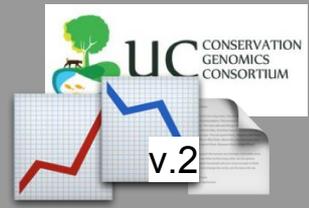
v.1

v.1

UNIVERSITY OF CALIFORNIA  
UCMERCED

SDSC  
SAN DIEGO SUPERCOMPUTER CENTER

INTERNET  
ARCHIVE



v.1

v.2

v.1

v.1

v.1



UC Curation Center



The Dawson Lab



v.1

v.1

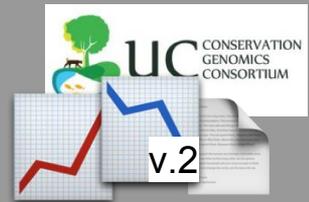
UNIVERSITY OF CALIFORNIA  
UCMERCED

SDSC  
SAN DIEGO SUPERCOMPUTER CENTER

v.1  
v.2



v.2  
v.1



v.1

v.2

Prototype  
status:



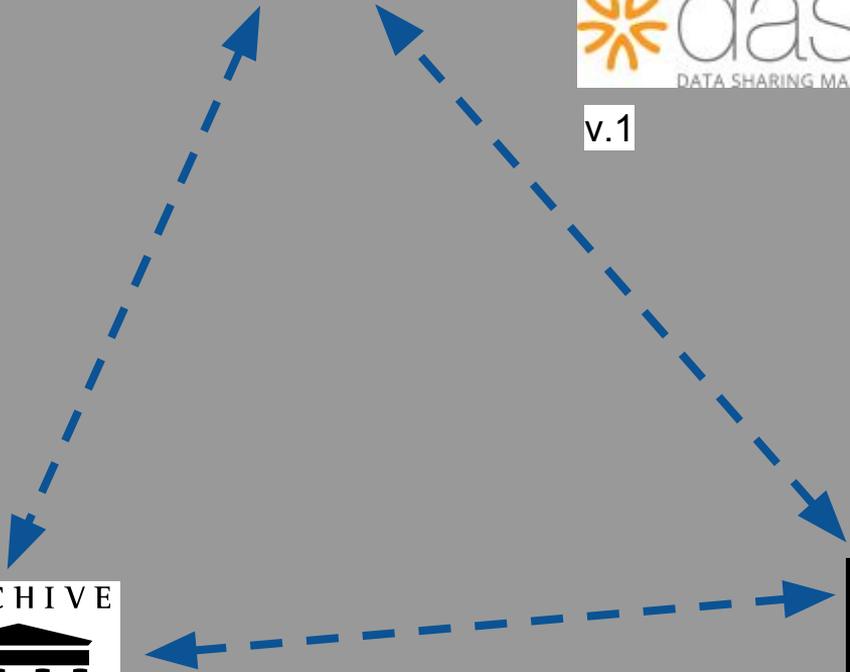
<1 TB  
Dat archive

v.1

v.1



v.1



Every institution contributes

Storage, bandwidth

Metadata on their collection

Commitment to preserve their collection

to the network

**CS&S**

@daniellecrobins

Any user can access

Information on collections

History of objects

Whole or partial datasets

from the network

Long term stewardship of  
scholarly work

Not a technology problem

Institutions + people +  
commitments



No technology can  
guarantee the long  
term stewardship of  
scholarly work



This project is about  
people

Existing infrastructure

+

Link trusted institutions

+

Develop community,  
handle uncertainty





Dat in the Lab stakeholders convening 2018

GORDON AND BETTY  
**MOORE**  
FOUNDATION



UNIVERSITY OF CALIFORNIA



**UC DAVIS**  
**LIBRARY**

**UCLA**

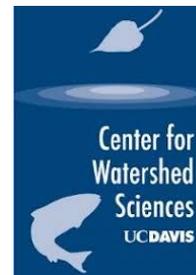


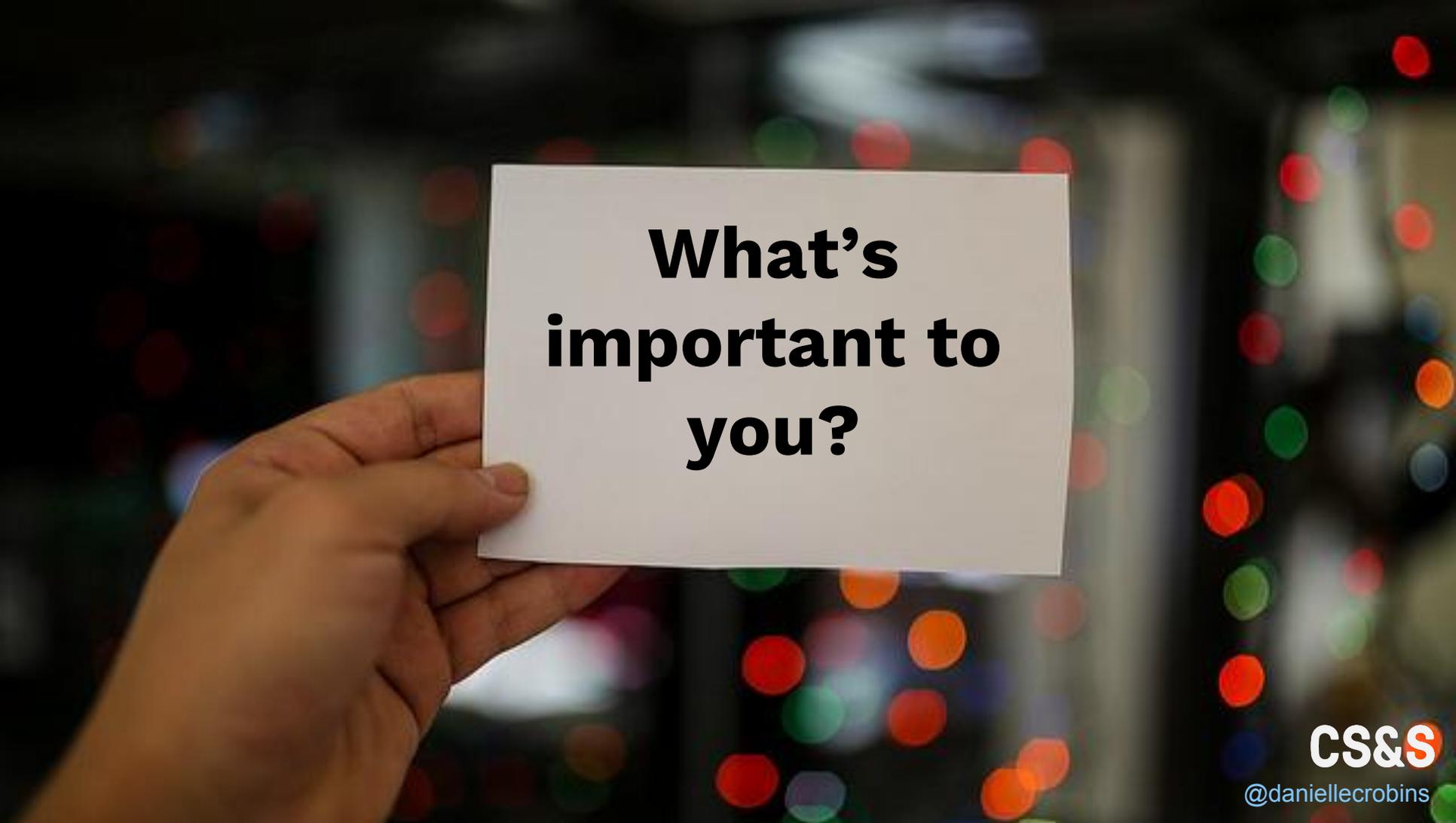
UC CONSERVATION GENOMICS CONSORTIUM



University of California  
**CDL**  
California Digital Library

BERKELEY  
**Library**  
UNIVERSITY OF CALIFORNIA



A hand is holding a white rectangular card in the center of the frame. The card has the text "What's important to you?" written on it in a bold, black, sans-serif font. The background is dark with many out-of-focus, colorful lights in shades of red, green, and orange, creating a bokeh effect. The lighting is soft, highlighting the hand and the card.

**What's  
important to  
you?**

**CS&S**

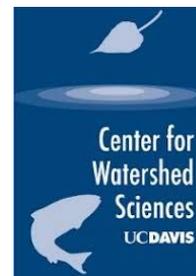
@daniellecrobins

# Thank you!

Extra thanks Dat in the Lab folks and the FORCE2018 organizing and program committees

DANIELLE ROBINSON, PhD  
Co-Executive Director  
[Code for Science & Society](#)

@daniellecrobins @codeforsociety



Reading & resources produced during the grant:

1. [Introductory blog](#)
2. Visits to [UC Davis](#) and [UC Merced](#) (blog)
3. [Dat in the Lab](#) (talk at SciData17) & [cartoon of talk](#)
4. [Self archiving with Dat](#) (blog)
5. [Data sharing between institutions](#) (blog)
6. [Virtual environments](#) (blog)
7. [A vision for decentralized data preservation across a network of libraries and trusted institutions](#) (talk at OnlineNorthwest)
8. [Challenges in implementing software containers](#) (blog)
9. [Decentralizing scholarly resources](#) (blog)
10. [Data Stewardship on the Decentralized Web](#) (talk and panel discussion at The Decentralized Web Summit)
11. The Dat Project, an open and decentralized research data tool (forthcoming paper in Scientific Data)
12. [A decentralized scholarly commons](#) (forthcoming talk at FORCE2018)
13. ANACAPA [repo](#), [README](#) and [toolkit](#)
14. CALeDNA Anacapa/CRUX Dat Container ([Linux/HPC](#) and [Windows/Mac Vagrant version](#))
15. [Anacapa Blog post](#)
16. ANACAPA paper, forthcoming

