




Expert Curation of SHARE Dataset



Jeffrey Spies, CTO and Co-founder, Center for Open Science
jeff@cos.io, [@jeffspies](https://twitter.com/jeffspies)

Cynthia Hudson-Vitale, Digital Data Outreach Librarian, Washington University in St. Louis
chudson@wustl.edu, [@cynhudson](https://twitter.com/cynhudson)

Judy Ruttenberg, Program Director, Association of Research Libraries
judy@arl.org, [@JudyRuttenberg](https://twitter.com/JudyRuttenberg)



Session overview

What is SHARE

Service Learning as Pedagogy

SHARE/LIS service learning project

SHARE Curation Associates

Future

Technology is important
to the modern librarian.

Technology continually evolves.



The modern librarian must continually learn in order to keep up with technology.

Recommendation:
Do not learn new technology.

Recommendation:
Learn how to learn new technology.

Recommendation:
Do not **teach** new technology.
Teach how to learn new technology.

In the real-world,
we (mostly) work independently.

In the real-world ,
we (mostly) ~~work~~ learn independently.

If we teach how to learn and
continue to support independent learning,
we give people resources.

In the real-world ,
we work on real projects.

If education is supposed to benefit
our work, and our work requires
real projects and continually learning, then
the world is our classroom.

Service learning balances **instruction** with a **pragmatic, real-life** opportunity to serve in the **community**.

The Pedagogy

Learning how to learn

...via big, meaningful, real-world projects that
are your projects.

“Big. Meaningful. Real-world. Mine.”

Learning is scary.
Learning is lonely.

Learning how to learn:
You are no longer under-resourced.

Learning is frustrating.
Frustration undermines learning.
This is a cyclic process.

But frustration is normal.
Make it work for you.

The role of frustration

Repeat the following:
Frustration is normal.
I am not different.

Frustration tells us
when we need help.

Frustration as a heuristic

Just starting: 45-60 minutes

Advanced: 90 minutes

Novel: 180 minutes

What doesn't work

Too many questions

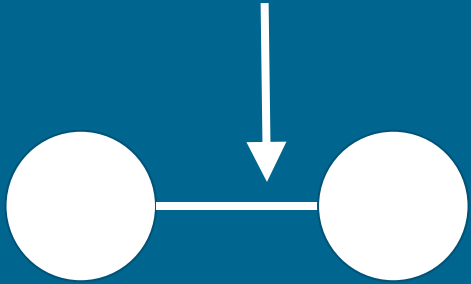
Not enough questions

Frustration tells us
when a problem is too big.

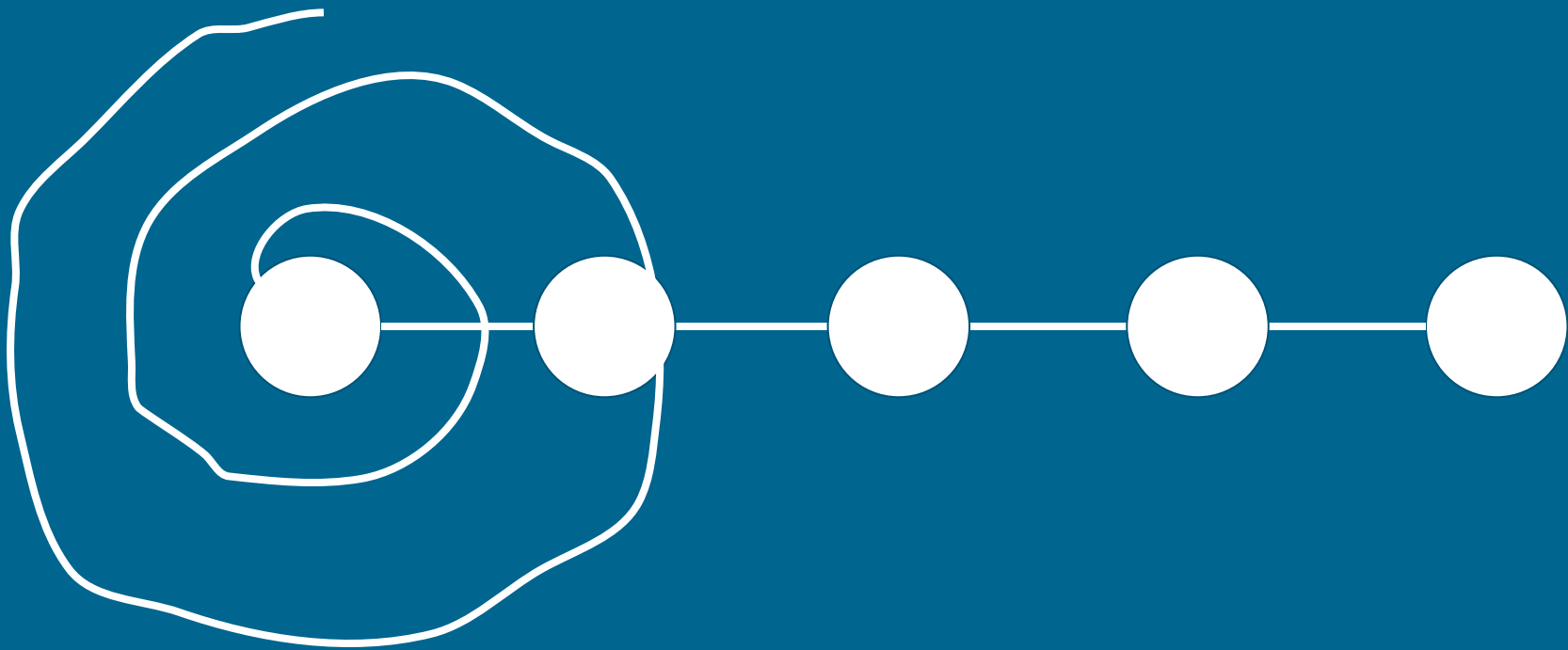


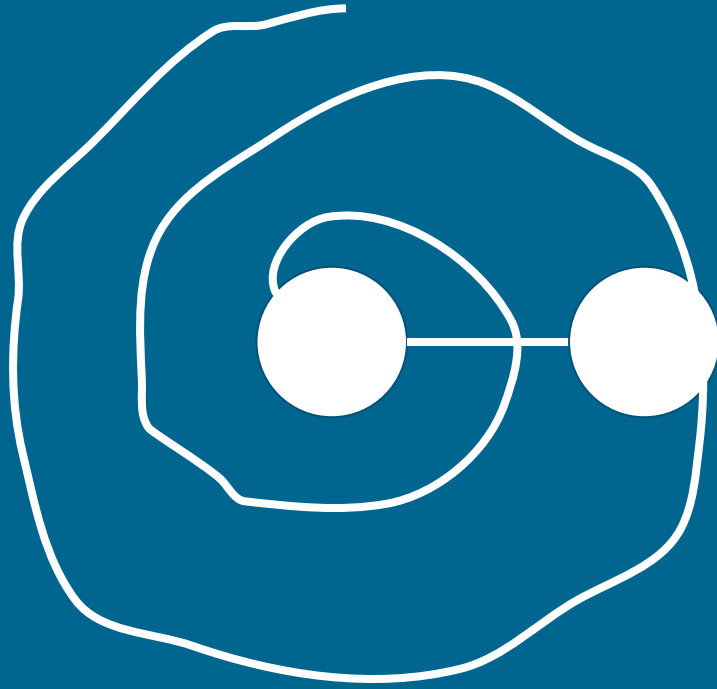


You are here.

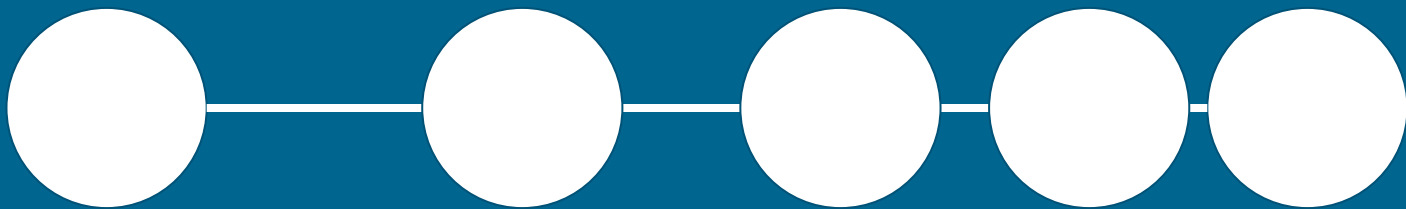


The problem is now consumable based upon where you are in your learning.





Spiraling around the solution--being frustrated--helped you solve the next problem.



Meaningful, real-world projects
foster diversity

Meaningful, real-world projects attract diversity

Those

- who care more the problem
- who care more about helping the community that needs the solution

than the technology used to solve the problem (but it still attracts them too)

To reinforce, frequently pitch:

Why is what you're working on the most important thing that **anyone** could be working on [for X | to solve Y]?

SHARE Community



LIS Graduate School Pilot: Background



SHARE is **gathering, cleaning, linking,** and **enhancing metadata** that describe research activities and outputs—from **data management plans** and **grant proposals** to **preprints, presentations, journal articles,** and **data repository deposits**—in order to make this metadata openly and freely available.

LIS Graduate School Pilot: Background

Digital curation is...

“the active management and enhancement of digital information assets for current and future use”

PREPARING THE WORKFORCE FOR DIGITAL CURATION

COMMITTEE ON FUTURE CAREER OPPORTUNITIES AND EDUCATIONAL REQUIREMENTS FOR
DIGITAL CURATION

BOARD ON RESEARCH DATA AND INFORMATION

Policy and Global Affairs

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

LIS Graduate School Pilot: The Goals

Provide student-centered, hands-on work that:
benefits the community-developed SHARE dataset,
equips individuals with transferrable skills in digital curation while building
technical confidence, and
improves local metadata about scholarly activities.



The Class

School of Information Science and Learning Technologies,
University of Missouri, Columbia

Heather Moulaison-Sandy, PhD

'Digital Libraries' graduate course

Spring 2016

Online

16 students

2 - 1.5 hour synchronous courses + homework



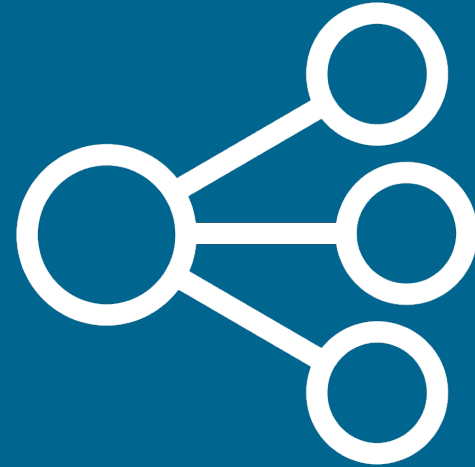
The Concepts & Skills

Application programming interface (API)

JSON

Metadata enhancements

Scholarly communications



Created by Emily van den Heever
from Noun Project

API/JSON

- Take 5 minutes
- Visit the SHARE schema site
 - Though there few required fields, there are many optional ones.
 - Choose one optional field, read the description , determine acceptable formats, and any requirements
 - Report back on one optional field you found and how you would use it

SHARE API

SHARE Search

4

access is not permitted.

supports both GET and POST requests. The data returned by both endpoints is JSON, by default

ments found,
s corresponding to the SHARE schema,
query took

parameters in common:

jtg/wiki/home/

API Query Syntax

API Query String

Below are four information requests your supervisor just gave you. Regroup in your virtual rooms and collaborate to write query strings that will most efficiently pull back the requested information. One person in your group will be responsible for reporting back.

1. Write an API search query string to search SHARE for source ASU and return records 2-30
2. Write an API query string to search SHARE and find all records by author Cynthia Traub
3. What is wrong with the query string below?
 1. [https://osf.io/api/v1/share/search/?q=\(shareProperties.source:wustlopendscholarship\)ANDfrom=5ANDsize=2](https://osf.io/api/v1/share/search/?q=(shareProperties.source:wustlopendscholarship)ANDfrom=5ANDsize=2)
4. Write an API query string to search SHARE for source WUSTL and find all records that are about biology

1. API documentation to
upon API's
ns and operators

2. 'fish'
sh'
combined with Boolean

`Properties.source:asu)AND(title:"fish")`

Search parameter 1

Boolean operator

Metadata Enhancements

80 WU institutional repository records

Each student enhanced 20 records

Quality assurance: Each enhancement validated 4 times

Validations checked by SHARE Ops team

Elements:

<freetoread>

<sponsorID>

<sponsorName>

<awardID>

<awardName>

<freetoread>

WASHINGTON UNIVERSITY
OPEN SCHOLARSHIP

Home About FAQ My Account

Enter search term: Search

in this series

Advanced Search

Notify me via email or RSS

Browse
Collections
Disciplines
Authors

Author Corner
Author FAQ

Home > ART_SCL > C_ART_SCL > BIO > BIO_FACTSBS > 38

BIOLOGY FACULTY PUBLICATIONS & PRESENTATIONS

Single Molecule Analysis of the Arabidopsis FRA1 Kinesin Shows that It Is a Functional Motor Protein with Unusually High Processivity

[Download](#)

Included in
Biochemistry, Cellmolec,
Biolsci, Commun, Plant
Biolsci, Commun

Author's Name
Chuanmei Zhu
Ran Dworkin, Washington University in St. Louis

Author's School
Arts & Sciences

Author's Department
Biology

Document Type
Article

Publication Date
9-2011

Originally Published In
Mol Plant. 2011 Sep;4(3):879-85. doi: 10.1093/mp/ksr077

SHERPA/RoMEO

... opening access to research

Home • Search • Journals • Publishers • FAQ • Suggest • About

English | Español | Magyar | Nederlands | Português

Publisher copyright policies & self-archiving

ISSNs Publisher names

Starts with contains ISSN

[Advanced Search](#) Search Reset

summary of permissions that are normally given as part of each publisher's copyright

RoMEO News [Blog](#) [Twitter](#) [More >>](#)

- [Dutch Version of RoMEO Released](#) - 22-Apr-2015
- [Nederlandse versie van RoMEO verschenen](#) - 22-Apr-2015
- [SHERPA/RoMEO API Version 2.4 has been Closed Down](#) - 26-Nov-2013

Special RoMEO Pages [More >>](#)

- [Publishers Allowing use of their PDFs in Repositories](#)
- [RoMEO Statistics](#)
- [Application Programmers' Interface \(API\)](#)

Additions and Updates [RSS1 Feed](#) [More >>](#)

- [Politechnika Śląska \(Silesian University of Technology\)](#) - Politechnika Śląska (Silesian University of Technology) - 05-Feb-2016
- [Mustafa Kemal Üniversitesi](#) - Mustafa Kemal Üniversitesi - 05-Feb-2016
- [Câmara Municipal de Lisboa, Arquivo Municipal](#) - Câmara Municipal de Lisboa, Arquivo Municipal - 05-Feb-2016

Other SHERPA Services

- [SHERPA/FACT](#) - Funders & Authors Compliance Tool
- [SHERPA/JULIET](#) - Research funders' open access policies

Jisc

<sponsor>

WASHINGTON UNIVERSITY
OPEN SCHOLARSHIP

[Home](#) [About](#) [FAQ](#) [My Account](#)

Enter search terms: Search

In this series

[Advanced Search](#)

[Notify me via email or RSS](#)

Browse

- [Collections](#)
- [Disciplines](#)
- [Authors](#)

Author Corner

[Author FAQ](#)

[Home](#) > [ART_SCI](#) > [C_ART_SCI](#) > [BIO](#) > [BIO_FACPUBS](#) > 78

BIOLOGY FACULTY PUBLICATIONS & I

Single Molecule Analysis of the Arabidopsis FRA1 Kinesin Shows that It Is a Functional Motor Protein with Unusually High Processivity

Chuanmei Zhu
[Ram Dixit](#), *Washington University in St. Louis* [Follow](#)

Author's School
Arts & Sciences

Author's Department
Biology

Document Type
Article

Publication Date
9-2011

Originally Published In
Mol Plant. 2011 Sep;4(5):879-85. doi: 10.1093/mp/ssp077

Tip on award data

- Look in the full text acknowledgements or search the full text document for 'funding' or 'grant'
 - awardName
 - awardIdentifier
 - sponsorName
- Imaging Sciences Pathway
5T90 DA022781
National Institute of Health

ACKNOWLEDGEMENTS

C.Z. was supported by grant 5T90 DA022781 from the imaging sciences pathway at Washington University. We thank the Pakrasi lab for access to the plate reader for the ATPase experiment. We thank Erica Fishel for critical reading of this manuscript and Amanda Tulin for her help with cloning FRA1.

FIGURE LEGENDS

Figure 1. Microtubule binding and ATPase activity of FRA1(707)-GFP. (A) The functional domains of FRA1 and FRA1(707)-GFP (top) and predicted coiled-coil domains of FRA1

What We Learned/Going Forward

Students feedback was positive

High quality metadata enhancements

Metadata is really really hard

Another pilot this summer

Scale this initiative to additional LIS programs

Expand skill development and technical competencies to library professionals

SHARE Curation Associates

July 2016–July 2017

Hoped for initial cohort of 20 library professionals

Skills

Information extraction through APIs and OAI-PMH feeds

Building content harvesters

Basic programming in Python

Methods and tools to automate data cleaning and metadata enhancement

SHARE Curation Associates

July 2016–July 2017

Strong response from library community—cohort will be ~35

Metadata, Digital Initiatives, Repository, Scholarly Communication, Archives

Responsibilities

Attend training and focus group session at COS in July

Conduct at least two additional trainings OR host a curate-a-thon

Spend 2–3 hours per week curating, linking, and/or enhancing the SHARE data set (and their own repo feed into SHARE)

SHARE Curation Associates



SHARE Curation Associates

Wide geographic diversity

Mix of public and private universities

80% are from ARL libraries

Diversity of technology skills and experience

Strategic directions

Contributes to workforce development

Links LIS curriculum and library workforce needs

Provides mentoring opportunities

Leverages new partnerships for new challenges

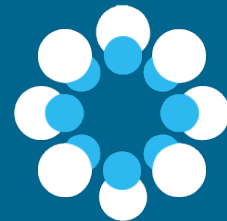
Coalesces the community around a major infrastructure project

There's a lot of work to do: get involved!

Contact us at

info@share-research.org

[@SHARE_research](#)



Find these slides on CNI's slide-share:

<https://osf.io/view/CNI2016/>

Extra Slides

