SHARE UPDATE

CNI Membership Meeting, 8 December 2014

Tyler Walters, SHARE Director and Dean of Libraries at Virginia Tech
Eric Celeste, SHARE Technical Director
Jeff Spies, Co-founder/CTO at Center for Open Science
WHO & WHAT IS SHARE?

SHARE is a higher education initiative to maximize research impact.
WHO & WHAT IS SHARE?

SHARE envisions an environment where researchers can keep interested parties seamlessly informed of their activities, where funders can easily determine the impact of their grant investments, and where institutions can readily collect and assess the output of their community members.
WHO & WHAT IS SHARE?

Advisory Board

Director, small operations group

Working Groups addressing key tasks
FUNDING

$1,000,000 to develop Notification Service and long term SHARE vision

March, 2014 through September, 2015
“Research universities are long-lived and are mission-driven to generate, make accessible, and preserve over time new knowledge and understanding.”

SHared Access Research Ecosystem, June 2013
MISSION

Infrastructure

Workflow

Policy
MISSION

Infrastructure

- Repository platforms/capacities
- Researcher networking platforms
- CRIS systems
- Standards and Protocols
- Identifiers

Workflow

- Multiple siloed systems = Administrative burden

Policy

- Public access
- Open access
- Copyright
- Data management & sharing
- Institutions’ internal information policies

SHARE
MISSION

Maximizing Research Impact

Infrastructure

Software (Notification Service and other components)
Open data and APIs
Encouraging standards

Workflow

New services to optimize communication; support research lifecycle

Policy

Best practices re: institutional policies
INSTITUTIONAL CONTEXT

US federal agencies join growing international trend to require public access to funded research

Measurable proliferation of institutional and disciplinary repositories

Premium on impact and visibility in HE
RESEARCH CONTEXT

- “Scholarly outcomes are contextualized by materials generated in the process and aftermath of scholarly inquiry.
- The research process generates materials covering methods employed, evidence used, and formative discussion.
- The research aftermath generates materials covering discussion, revision, and reuse of scholarly outcomes.”

(Lavoie, et al, OCLC Research, 2014)
RESEARCH CONTEXT

- “Scholarly outcomes are contextualized by materials generated in the process and aftermath of scholarly inquiry.

- The research process generates materials covering methods employed, evidence used, and formative discussion.

- The research aftermath generates materials covering discussion, revision, and reuse of scholarly outcomes.”

(Lavoie, et al, OCLC Research, 2014)
RESEARCH CONTEXT

- “Scholarly outcomes are contextualized by materials generated in the process and aftermath of scholarly inquiry.
- The research process generates materials covering methods employed, evidence used, and formative discussion.
- The research aftermath generates materials covering discussion, revision, and reuse of scholarly outcomes.”

(Lavoie, et al, OCLC Research, 2014)
RESEARCH LIBRARIES

↑ collaboration among institutions shift from collections as *products* to collections as *components* of the academy’s knowledge resources.

library is supporting and embedded within the processes of scholarship.
STATEMENTS OF SUPPORT
NOTIFICATION SERVICE

Knowing who is producing what, and under whose auspices, is critical to a wide range of stakeholders—funders, sponsored research offices, government agencies, tenure and promotion committees, repository managers, and the research community.
CONSUMERS OF RESEARCH RELEASE EVENTS

Campus Repositories

Sponsored Research Offices

Funders
CENTER FOR OPEN SCIENCE

“We foster openness, integrity, and reproducibility of scientific research”

http://cos.io & http://osf.io
Modular
Reliable
Flexible
Scalable
Gather''

Notify'

Providers

Consumers
Application Programming Interface
Push API

Content Harvesters
Gather

Notify

Providers

Consumers
Provider: stcloudstate  

11:05
Provider: stcloudstate

11:05

GET
stcloud.edu/api/
Provider: stcloudstate

11:05

Queue

- GET plos.org/api/
- GET stcloud.edu/api/
- GET columbia.edu/api/
- GET mit.edu/api/
- GET vtech.edu/api/
- GET dataONE.org/api/
Provider: stcloudstate

11:05

Queue

GET plos.org/api/
GET stcloud.edu/api/
GET columbia.edu/api/
GET mit.edu/api/
GET vtech.edu/api/
GET dataONE.org/api/

worker
Request

GET
http://repository.stcloudstate.edu/do/oai/?

verb=ListRecords
&metadataPrefix=oai_dc
&from=2014-10-02T00:00:00Z

Response

<record>
  <date> … </date>
  <title> … </title>
  <contributors>
    <name>…</name>
    <name>…</name>
  </contributors>
  <description>
    …
  </description>
  <id>…</id>
</record>
GET stcloud.edu/api/

- Archive Request
- Archive Response
- Archive Data
GET
stcloud.edu/api/

- Archive Request
- Archive Response
- Archive Data
- Normalize
GET
stcloud.edu/api/

- Archive Request
- Archive Response
- Archive Data
- Normalize
- Send to OSF
Notify

Consumers
- RSS/Atom
- PubSubHubbub
- ResourceSync
- OSF Search Engine
- OSF API
/rss/?q=open AND science AND source:scitech
CURRENT STATE

24 source; 251,596 events
NEXT STEPS

Push protocol
Recruit push providers

Consumption of notifications
Provide subscription methods
Recruit trial subscribers

Public release
Early 2015 beta release
Fall 2015 first full release
SOME EARLY LESSONS

Metadata rights issues.
Some sites not sure about their right to, for example, share abstracts.

Metadata inclusion and consistency.
Most of our sources do not even collect email addresses of authors, much less universal identifiers such as ORCID or ISNI. Most sources make no effort to collect funding information or grant award numbers. This data needs to be collected and distributed to make effective notifications.

The need for a Phase II.
Some consumers will want the enhanced records it will provide.
METADATA RIGHTS

Does metadata gathering violate your terms of service?  
If so, are we granted explicit, written rights to gather data?

Does metadata gathering violate your privacy policy?  
If so, are we granted explicit, written rights to gather data?

Does our sharing the metadata we gather from you violate your policies?  
If so, are we granted explicit, written license to share the metadata?

Do you use an explicit license for your metadata (for example, CC Zero)?  
If not, do you have plans to explicitly license the content?
VARIETY AND AVAILABILITY

• We accept that we will have a variety of providers with a variety of expressions.

• But we need some key identifiers to be available in order to create effective notifications.
GUIDELINES?

• See https://www.coar-repositories.org
INCLUSION OF IDENTIFIERS

• Researcher identifiers such as ORCID, ISNI, and so on.
• Funding identifier such as FundRef.
• Grant award identifiers.
• Further metadata elements encouraged by COAR, CASRAI and others.
CONSISTENCY ACROSS PROVIDERS

• We can manage the variety.
  ...however...
• Consistency reduces errors.
• Consistency simplifies preparing for new providers.
• Consistency will be required for push reporting.
For Systems via Protocol & API

SHARE Notification Service → SHARE Registry → SHARE Discovery

For People

timely, structured, comprehensive

organized and related source of linked data

searchable and friendly
CHALLENGES

• Adoption of key identifiers just getting underway, requires international collaboration,
• Inferences prone to error,
• Duplicate detection difficult,
• Scale quite large, not well understood,
• This is a never-ending task requiring sustainable funding and governance.
SHARE Notification Service including Reconciliation?

For Systems via Protocol & API

SHARE Discovery

For People

timely, structured, comprehensive, constantly adjusting to new input

searchable and friendly
PHASE II BENEFITS

• Researchers can keep everyone informed by keeping anyone informed,
• Institutions can assemble more comprehensive record of impact,
• Open access advocates can hold publishers accountable for promises,
• Other systems can count on consistency of metadata from SHARE.
LOOKING FORWARD

Business planning: agile governance and institutional sustainability

Strengthening international partnerships

Workflow pilots and prototypes with several institutions
LOOKING FORWARD

Identifiers and platforms for new, hybrid, and more granular forms of research output

Contributor roles & other research administration data

Higher education policies on non-exclusive copyrights
CONTACT US

www.arl.org/share

bit.ly/sharegithub

share@arl.org

www.facebook.com/SHARE.research

www.twitter.com/share_research