

OPEN REPOSITORIES 2023

# The Harvard Data Commons: Leveraging COAR Notify to create interoperability between Harvard data and institutional repositories

**Colin Lukens**

*Harvard Library Open Scholarship and Research Data Services*

**Andrew Woods**

*Harvard Library Technical Services*

June 14, 2023

<https://bit.ly/or23-hdc>



# Harvard Library - Advancing Open Knowledge

*“We will advance a new model of a distributed repository network and publishing services that exemplify collaborative, equitable approaches to scholarly communications and enable seamless access to a range of digital content, such as research data as well as open access publications.”*

[Advancing Open Knowledge](#)

# Harvard Dataverse Overview

Harvard's open online repository for sharing, preserving, citing, exploring and analyzing research data.

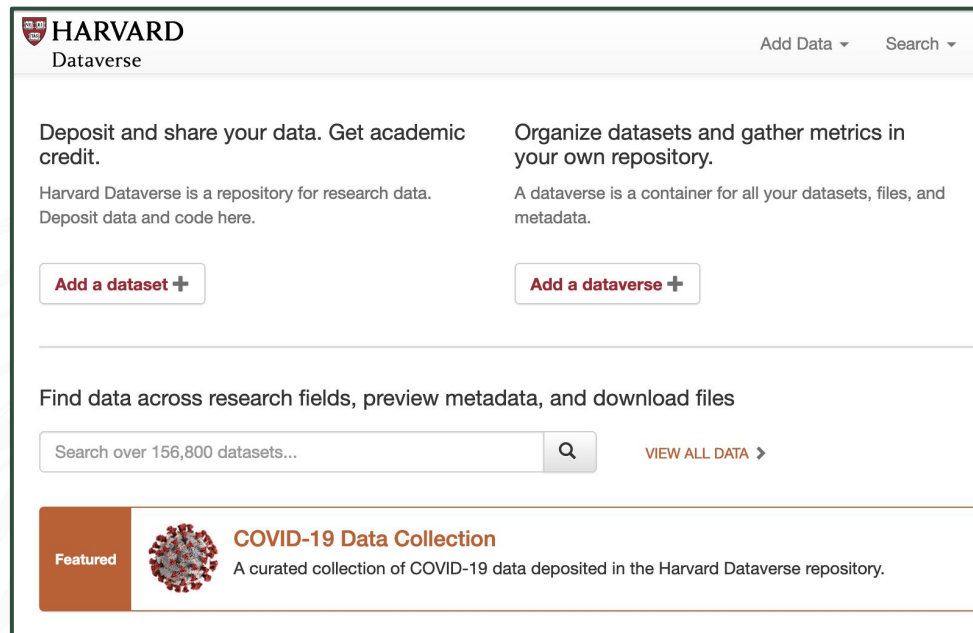
Over 1.5M downloads for 2022-2023

Over 50M downloads overall

Over 6,000 published collections

Over 164,000 datasets

Over 2M files

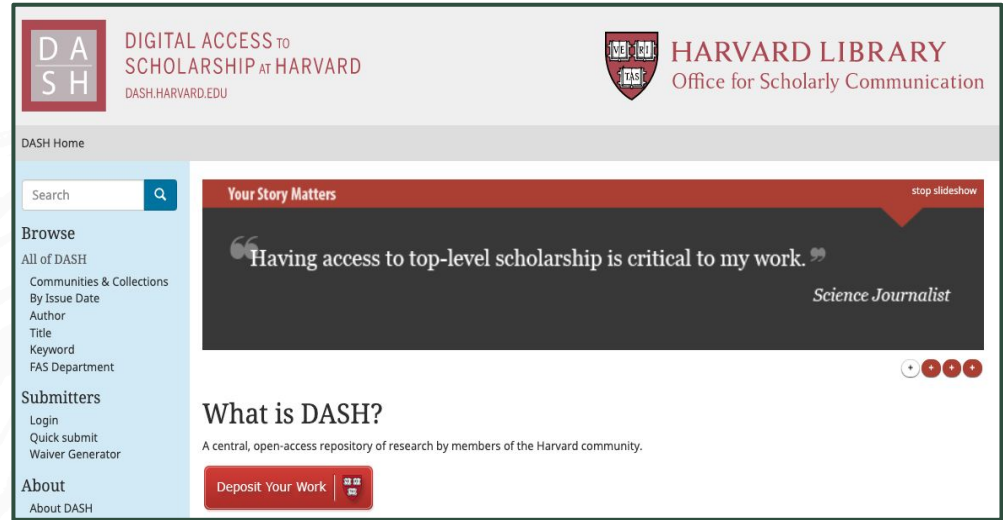


<https://dataverse.harvard.edu/>

# DASH Overview

Harvard's central, open-access, institutional repository, providing a global audience nearly 58,000 works of university-based scholarship. DASH, launched in 2008, now sees an average of 7M downloads per year.

Built on DSpace



<https://dash.harvard.edu/>

# COVID-19 data and publications

## Crosslinking COVID resources in Harvard's DASH and Harvard Dataverse repository



**DASH** DIGITAL ACCESS to SCHOLARSHIP at HARVARD  
DASH.HARVARD.EDU

**HARVARD LIBRARY** Office for Scholarly Communication

[DASH Home](#) »

Search

**Browse**  
All of DASH  
Communities & Collections  
By Issue Date  
Author  
Title  
Keyword  
FAS Department

**Submitters**  
Login  
Quick submit  
Waiver Generator

**About**  
About DASH  
DASH Stories  
DASH FAQs  
Accessibility  
COVID-related Research  
Terms of Use  
Privacy Policy

### Calling on Harvard's coronavirus researchers

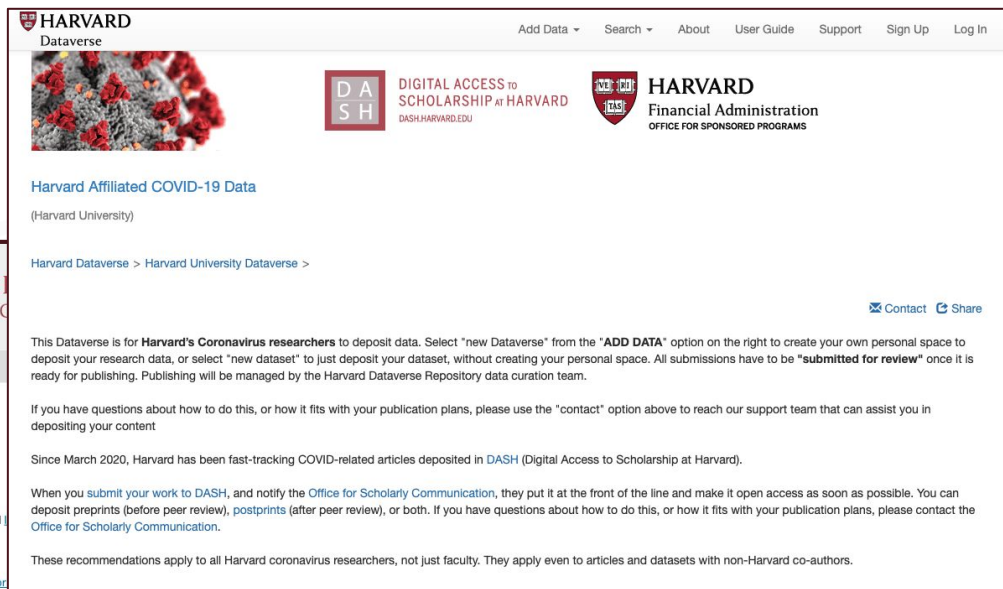
When you have COVID-related research ready to share, we recommend two steps.

1. [Deposit](#) your article(s) in [DASH](#) (Digital Access to Scholarship at Harvard).
  - Since March 2020, we've been fast-tracking COVID-related research into DASH. When you [submit your work](#), and the front of the line and make it open access as soon as possible.
  - You can deposit preprints (before peer review), [postprints](#) (after peer review), or both.
  - If you have questions about how to do this, or how it fits with your publication plans, please contact the [Office for Scholarly Communication](#).
2. [Deposit](#) your dataset(s) in the [COVID-19 Harvard Dataverse repository](#).
  - If you have questions about how to do this, or how it fits with your publication plans, please contact [Harvard Dataverse support](#).

These recommendations apply to all Harvard coronavirus researchers, not just faculty. They apply even to articles and datasets with non-Harvard co-authors.

Thanks for doing this important research and thanks for sharing it with the world!

[Harvard Office for Scholarly Communication](#) (for DASH)  
[Harvard Institute for Quantitative Social Science](#) (for Dataverse)



**HARVARD Dataverse**

Add Data Search About User Guide Support Sign Up Log In

**DASH** DIGITAL ACCESS to SCHOLARSHIP at HARVARD  
DASH.HARVARD.EDU

**HARVARD** Financial Administration  
OFFICE FOR SPONSORED PROGRAMS

### Harvard Affiliated COVID-19 Data

(Harvard University)

[Harvard Dataverse](#) > [Harvard University Dataverse](#) >

[Contact](#) [Share](#)

This Dataverse is for **Harvard's Coronavirus researchers** to deposit data. Select "new Dataverse" from the "ADD DATA" option on the right to create your own personal space to deposit your research data, or select "new dataset" to just deposit your dataset, without creating your personal space. All submissions have to be **"submitted for review"** once it is ready for publishing. Publishing will be managed by the Harvard Dataverse Repository data curation team.

If you have questions about how to do this, or how it fits with your publication plans, please use the "contact" option above to reach our support team that can assist you in depositing your content.

Since March 2020, Harvard has been fast-tracking COVID-related articles deposited in [DASH](#) (Digital Access to Scholarship at Harvard).

When you [submit your work to DASH](#), and notify the [Office for Scholarly Communication](#), they put it at the front of the line and make it open access as soon as possible. You can deposit preprints (before peer review), [postprints](#) (after peer review), or both. If you have questions about how to do this, or how it fits with your publication plans, please contact the [Office for Scholarly Communication](#).

These recommendations apply to all Harvard coronavirus researchers, not just faculty. They apply even to articles and datasets with non-Harvard co-authors.

COVID-19 works in DASH (Mar-Sept 2020): ~40 works; 437K downloads

277 related COVID-19 files in Harvard Dataverse Repository

# Notify for DASH and Harvard Dataverse repository?



Created by Deuxamis moon  
from the Noun Project



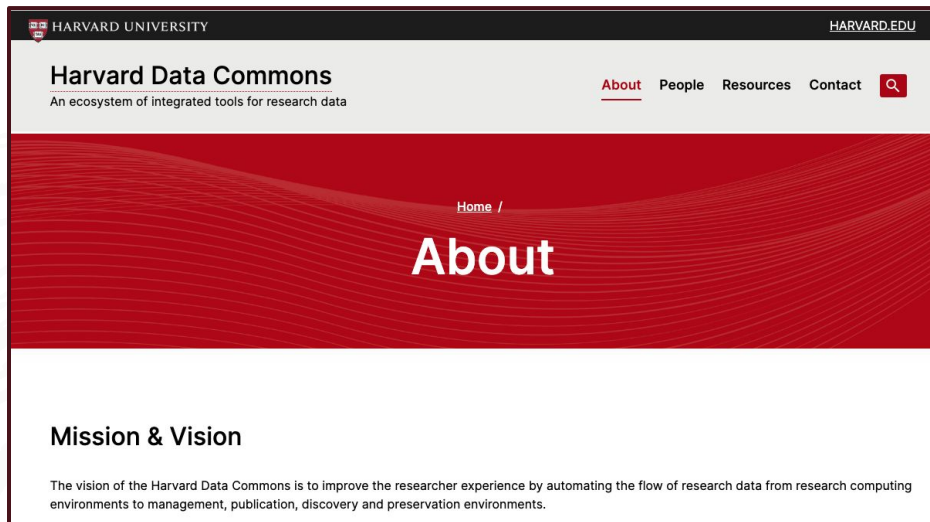
Created by Steve Morris  
from the Noun Project



# Harvard Data Commons

**Vision** of the Harvard Data Commons:

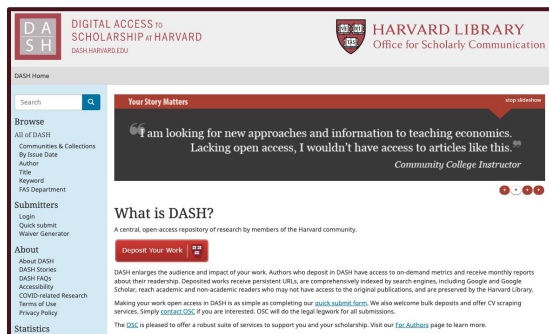
*To improve the researcher experience by automating the flow of research data from research computing environments to management, publication, discovery and preservation environments.*



<https://sites.harvard.edu/harvard-data-commons/>



# HDC Objective 3B - COAR:Notify for DASH and Harvard Dataverse repository



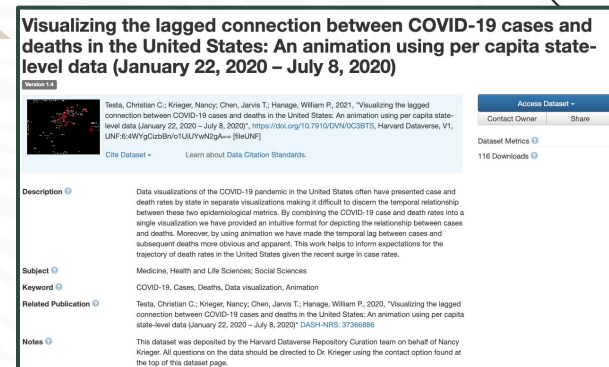
Created by Deuxamis\_moon  
from the Noun Project



Created by Steve Morris  
from the Noun Project



Created by Steve Morris  
from the Noun Project





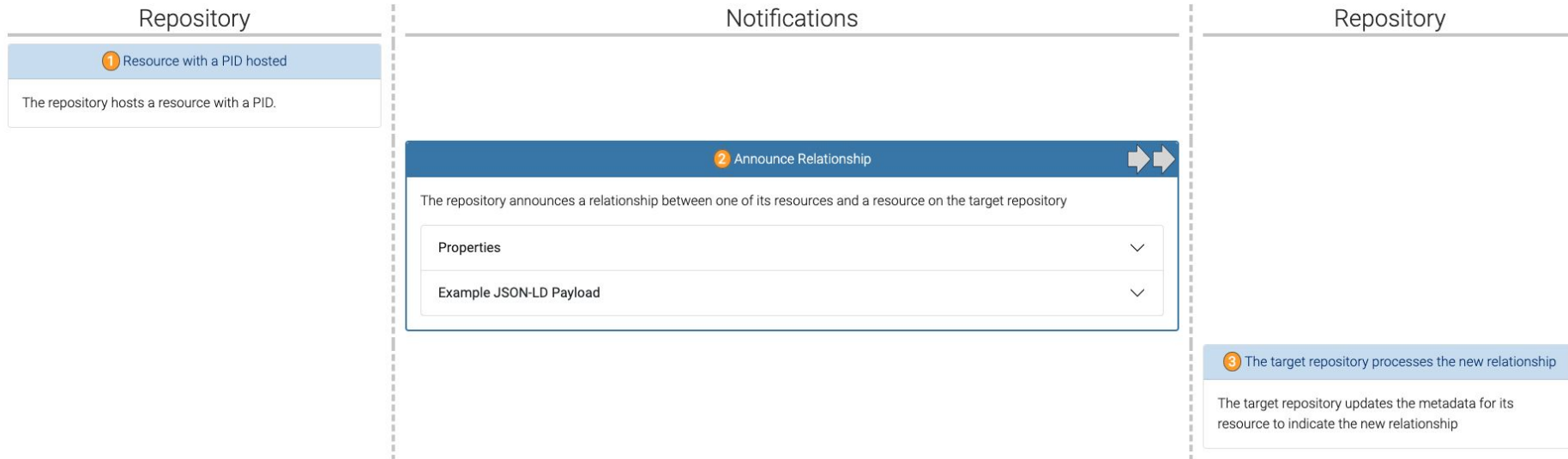
[Home](#) » [Example Scenarios](#) » Repository announces relationship between local resource and second resource hosted on another repository

## Example Scenario

# Repository announces relationship between local resource and second resource hosted on another repository

See these use-cases to understand this scenario in context: [Dissemination](#)

A repository announces a relationship between two resources, one of which is hosted by the repository, the other by the target system (which could be another repository)



## Notification Pattern

## Announce Relationship

**Description:** This pattern is used to announce a relationship between two resources

**Pattern Categories:** [Announcements](#)

### Properties

#### @context

The **@context** must include:

- <https://www.w3.org/ns/activitystreams>
- <https://purl.org/coar/notify>

#### id

Notify payloads must describe an AS 2.0 **activity**. The **activity** has an **id** which must be a URI, and the use of URN:UUID is recommended. An HTTP URI may be used instead, but in such cases the URI should resolve to a useful resource.

#### type

The **activity** has a **type** which should include *Announce & coar-notify:RelationshipAction*

#### actor

The **actor** describes the party responsible for this activity. The **actor** has:

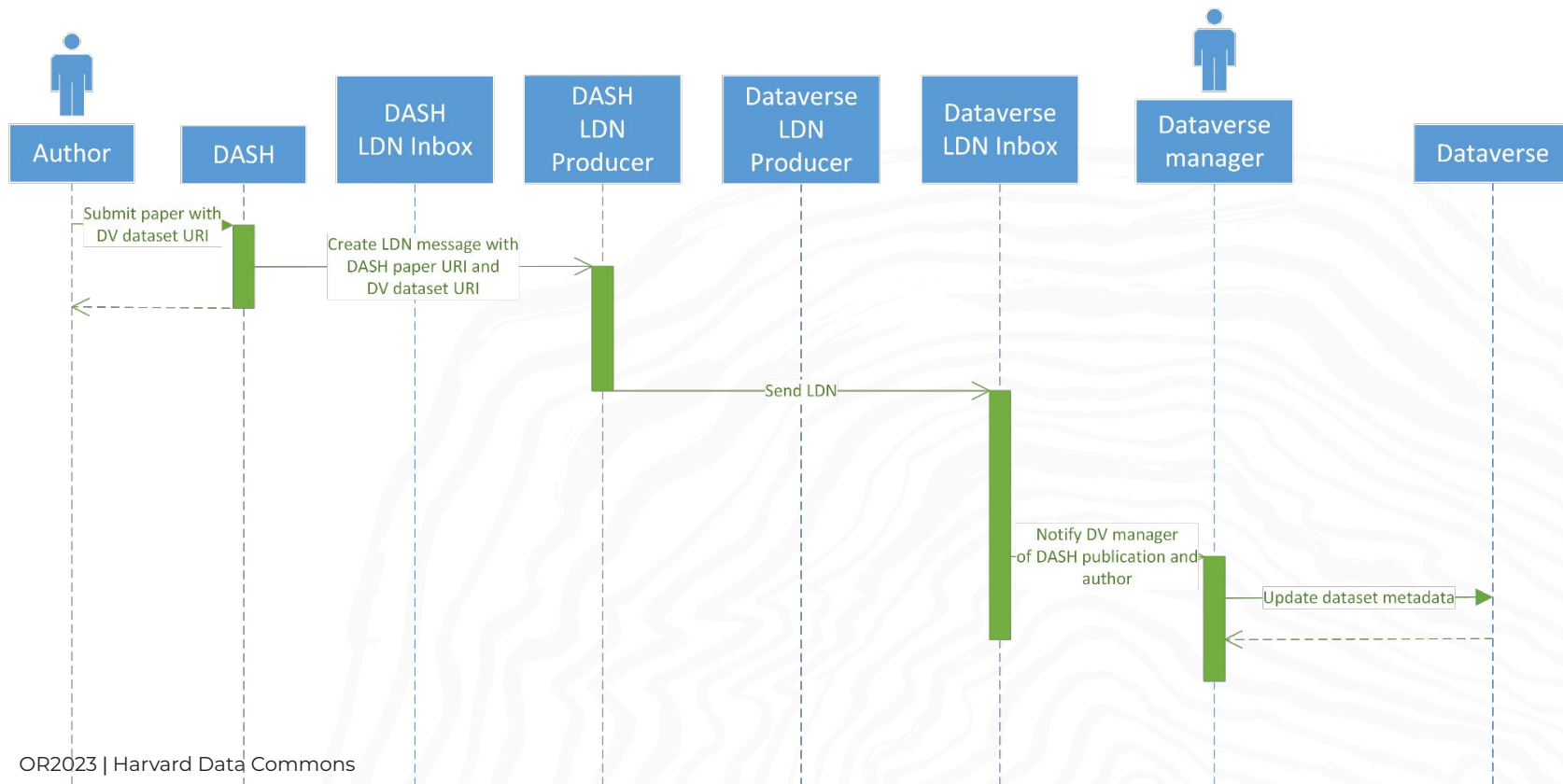
- An **id** which must be a URI (identifies the actor) (HTTP URIs are preferred, but any valid URI is accepted)

### Example JSON-LD Payload

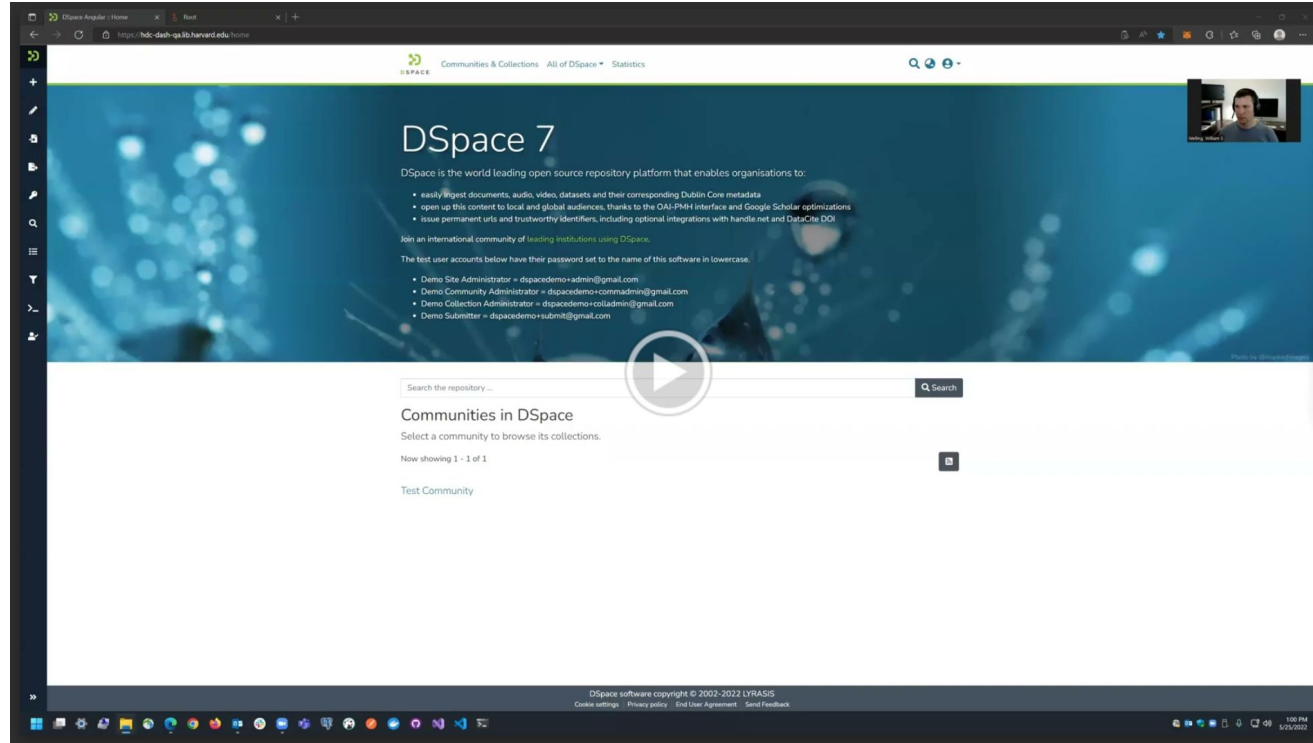
```
{
  "@context": [
    "https://www.w3.org/ns/activitystreams",
    "https://purl.org/coar/notify"
  ],
  "actor": {
    "id": "https://research-organisation.org",
    "name": "Research Organisation",
    "type": "Organization"
  },
  "context": {
    "id": "https://another-research-organisation.org/repository/datasets/item/201203421/",
    "ietf:cite-as": "https://doi.org/10.5555/999555666",
    "type": "sorg:AboutPage",
    "url": {
      "id": "https://another-research-organisation.org/repository/datasets/item/201203421/data_archive.zip",
      "mediaType": "application/zip",
      "type": [
        "Article",
        "sorg:Dataset"
      ]
    }
  }
}
```

<https://notify.coar-repositories.org/patterns/announce-relationship/>

Deposit to DASH with  
URI of Dataverse dataset



# In Action



<https://bit.ly/or23-hdc-demo>

# Status and Next Steps

## Dataverse

- COAR Notify features currently in 'main' project

## DSpace / DASH

- Prototype of COAR Notify in local DSpace7
- Implement COAR Notify in DSpace7 (8?)

## Collaboration

- Additional COAR Notify use cases: overlay journals?

# Acknowledgements

Sonia Barbosa

Jim Myers

Ceilyn Boyd

Tania Schlatter

Colleen Cressman

Stu Snyderman

Simone Biver-LeBlanc

Krista Valladares

Gustavo Durand

William Welling

Ardys Kozbial

Len Wisniewski

With thanks and gratitude to the COAR Notify Project team for inspiring this work:  
Kathleen Shearer, Paul Walk, Martin Klein, Eloy Rodrigues, Herbert Van de Sompel,  
Patrick Hochstenbach

# Thank You

**Colin Lukens**

colin\_lukens@harvard.edu

**Andrew Woods**

andrew\_woods@harvard.edu