Open Scholarship in the Humanities
Support at the UVa Library

SHERRY LAKE

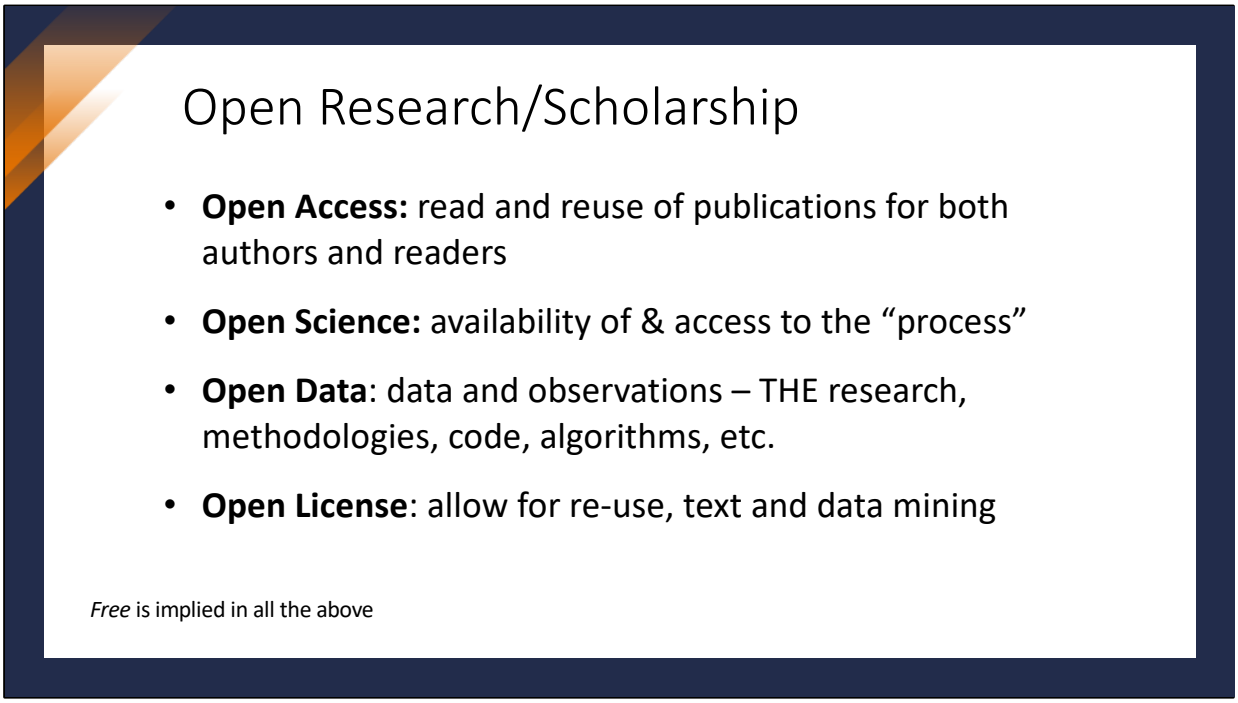
SCHOLARLY REPOSITORY LIBRARIAN

UNIVERSITY of VIRGINIA
LIBRARY

September 27, 2022

Hello. I'm Sherry Lake, Scholarly Repository Librarian at the University of Virginia
I want to thank Center for Digital Editing, the University of Virginia - Digital Publishing Cooperative, and eLaboratories for giving me the chance to talk about Open Scholarship in the Humanities and how we support Openness at the UVa Library.

This presentation will introduce "Open Access" and what it means in the humanities, highlighting differences between the humanities and the science disciplines.
It will conclude with an overview of the Library's supporting infrastructure - UVa's Institutional Repository Libra



Open scholarship (also referred to as "open science" and "open research") is the idea that, to advance knowledge, research results of all kinds should be openly shared as early as is practical. Open scholarship encompasses all disciplines, including science, the professions, arts and the humanities

As an element of open scholarship, open access is the ability to freely read and reuse publications. Being able to freely share published work which gives free access(removes price barriers to access) to ALL readers.

Open science in particularly aims to ensure the free availability and usability of scholarly publications, the data that result from scholarly research, and the methodologies, including code or algorithms, that were used to generate those data.

To truly be open, licenses should allow for re-use, including text and data mining - Free to re-use - removing permission barriers

copyright and licensing restrictions that require permission before one may copy, download, store, redistribute, crawl, or link to the texts
Licenses:
To avoid "copyright" legal restrictions – for Open Scholarship it is recommended that data have license with the least amount of restrictions. Think about who might benefit for re-use of your "content" – blogs, papers, ettc.
[CC - ND or NC avoid..... Do CC-BY or better CC-0]

Open access and free use (free from legal) can only be granted by content owner. If there are co-authors, you should makes sure all agree.

The next slides I'll discuss Open Data and the Humanities.

Data in Arts & Humanities

- Some data will be personal, and may not be factual
  - Content not necessarily important, but the presentation, arrangement... The expression
- Higher incidence of non-digital materials
- Digital data from Humanities is likely an outcome of the creative research process as opposed to an input to a workflow

Humanities Data: A Necessary Contradiction – Miriam Posner's Blog. (2015). https://miriamposner.com/blog/humanities-data-a-necessary-contradiction/. Accessed 14 Sept. 2021.

So if you ask a humanities scholar, would they say they have "data"? Maybe not because they use more precise terminology - primary sources, secondary sources, documents, bibliographies, annotations, notes, etc

For starters, Humanists don't always think of their sources or influences as 'data'.
When you call something data, you imply that it exists in discrete, fungible units; that it is computationally tractable; that its meaningful qualities can be enumerated in a finite list; that someone else performing the same operations on the same data will come up with the same results. This is not how humanists think of the material they work with.

few traditional humanists would call their source material "data" – so how would you describe what you use to do your research?
Humanists "engage with evidence" - research assets

Even if you have datasets (databases of old text, *mining* data from historical documents, sales of old art work) and do "analysis" it still
requires a lot of interpretive work.

Data in Arts & Humanities

Sharing Challenges

- Identifying “humanities data”
- Integrating 3rd party copyright sources
- Converting Formats (if possible)
- Lack of humanities data sharing policies (requirements)

So how would you identify your “data” – think about all materials and assets you collect, generate and use during all stages of the research lifecycle.

Think of all your research assets as research data that could be potentially reused by other scholars.

The …Lists, tables, spreadsheets, and "objects" physical and digital - photos, documents/manuscripts, sound, video, These aren’t so hard to “share” as long as you have copyright to do so. That’s another problem with sharing humanities work – considering how to include 3rd party copyright sources

Another problem is sharing your “project” webpage – interactive output. Need to think about how to share. What is more important; the experience of the webpage, or is it the content?

Unlike the sciences where there are many funder mandates to share (and have been for years), the Lack of polices for sharing in the humanities, makes these scholars have less experience with sharing

Rebecca Grant, F1000 Poster OASPA

Grant, Rebecca. *Opening up Humanities Scholarship: Open Humanities Data on the Routledge Open Research Publishing Platform*. https://oaspa.org/wp-content/uploads/2022/09/ROR-OASPA-2022-Poster-FINAL-Rebecca-Grant.pdf. OASPA Online Conference on Open Access Scholarly Publishing 2022.

Data in Arts & Humanities

What To Share

- all materials and assets scholars collect, generate and use during all stages of the research cycle, such as:
  - new data - field notes, survey responses, interview transcripts
  - Reused existing datasets
  - Other source materials – from libraries, archives, museums

So to think about "sharing" making your data open…

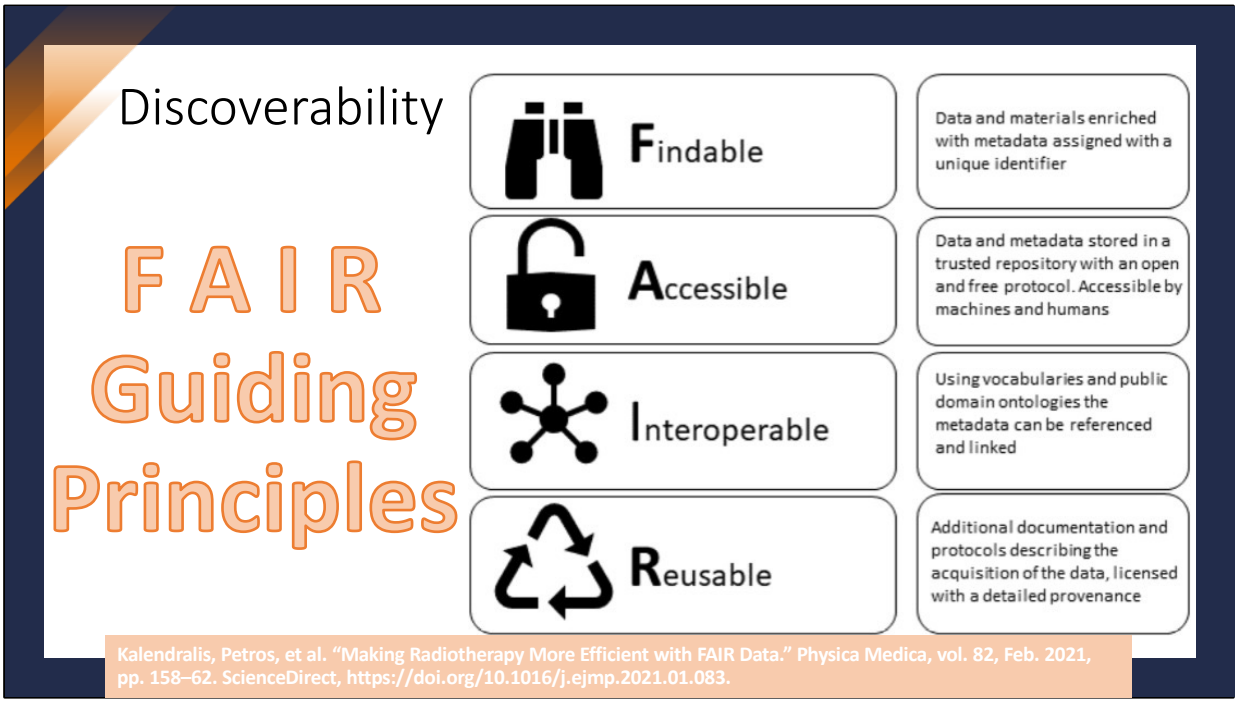
The data and sources underpinning your article may consist of archival documents, museum objects, audio files, AV material, images, transcripts, field notes, survey results, socioeconomic-, qualitative and quantitative datasets or other materials, which formed the input to, or output of your research. They may consist of physical objects, digitized materials or born digital files. Depending on your study design, you may have reused existing data or sources, or generated new datasets. Photos, corpus of text

Again, Think of all your research assets as research data that could be potentially reused by other scholars. Consider how useful it would be for your own work if others shared their data.

provide access to all of the data and sources you have generated or reused in your research. This is a key step to ensure that your research and methods (if you use any analysis) are transparent and that your results can be reproduced (where relevant).

•If you generated new datasets (for example field notes, survey responses, interview transcripts), you must deposit them into an appropriate data repository, and describe how they can be accessed and reused by others in your Data Availability Statement.
•If you reused existing datasets (for example Open Government datasets or data generated by another researcher) you must describe how they can be accessed and reused by others in your Data Availability Statement.
•If you used other source materials, for example from libraries, archives or museum collections, you

must describe how they can be accessed and reused by others in your Data Availability Statement. You should assume that if you have written a research article, there are data or source materials associated with your study, and you should therefore provide readers of your article with enough information to allow them to be accessed and potentially reused.



How – where to share: You may have heard about the FAIR Data Principles. Or share your "data" FAIRly.

The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*

The principles were formed to facilitate knowledge discovery by assisting humans and machines in

- discovery of
- access to
- integration and analysis of data and their associated algorithms and workflows.

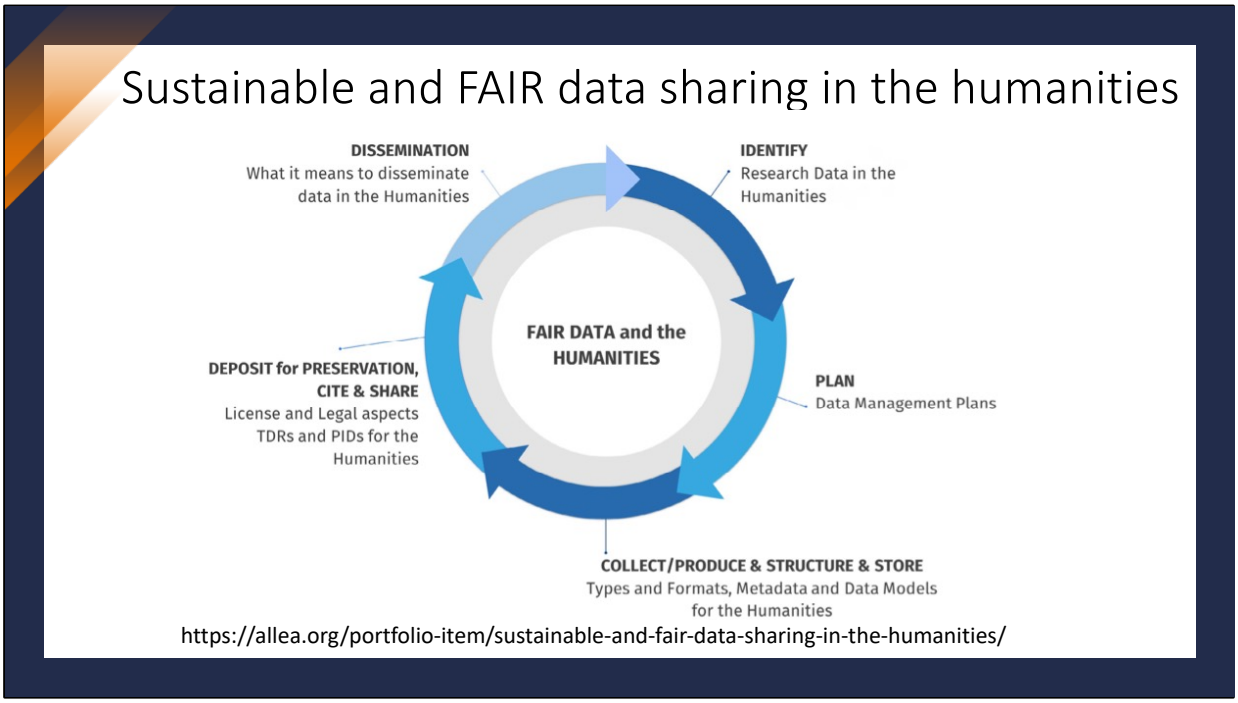
Making data (and metadata):

- Findable - *To be **findable (F)** or discoverable, metadata and data should be easy to find for both humans and computers* **[DOI assigned]**
- Accessible - *To be broadly **accessible (A)** data and metadata should be retrievable in a variety of formats that are sensible to humans and machines using persistent identifiers*
- Interoperable - *To be **interoperable (I)** use an open, well defined vocabulary. the data need to interoperate with applications or workflows for analysis, storage, and processing.*
- Reusable - *To be **reusable (R)** use should be easy*

- ❑ a clear and accessible data usage license
- ❑ (meta)data are associated with their provenance (documentation & contextual info cite when able, and describe how to access the sources, where they are located)

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016, March 15). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data,* 3, 160018. https://doi.org/10.1038/sdata.2016.18

Graphic: Kalendralis, Petros, et al. "Making Radiotherapy More Efficient with FAIR Data." Physica Medica, vol. 82, Feb. 2021, pp. 158–62. ScienceDirect, https://doi.org/10.1016/j.ejmp.2021.01.083.



https://allea.org/portfolio-item/sustainable-and-fair-data-sharing-in-the-humanities/
Copy of the report: https://virginia.box.com/s/ej6v5in5018rssbmepwy3ujvfbk9ho5i

ALLEA presents recommendations for sustainable data sharing in the humanities

Many research organisations working to both refine and expand how the principles can be adopted by scholarly communication practices. Research funders are incorporating FAIR data practices into their guidelines,

FAIR becoming influential beyond research - used as principles for managing and providing access to cultural heritage data

This is a report from All European Academies
recommendations to make digital data in the humanities "Findable, Accessible, Interoperable and Reusable", in line with the FAIR principles. The document is designed as a practical guide on how humanities could harness digital approaches and data management processes in ways to enhance scholarship and ensure that research outputs are sustainable over the long term.

This report provides recommendations to humanities scholars with the understanding that the humanities themselves are diverse and data practices and demands vary from scientific data research.

Presents data management as lifecycle - an ecosystem of research exchange & knowledge building the outputs of one research "part" becoming the inputs to another.

How pre-prints help humanities researchers

- sharing ideas and encouraging feedback and collaboration early on in the research process makes research stronger

- new research ideas that are not yet developed into their full, final form
  - other scholars offer constructive feedback
  - extended articles can be developed

Edwards, Caroline. 2014. "How Can Existing Open Access Models Work for Humanities and Social Science Research?". *Insights* 27 (1): 17–24.
DOI: http://doi.org/10.1629/2048-7754.135

Another way to "share" your research is with pre-prints….

Some "simple" ways of opening your research or parts of your research
Pre-print is a version of your article before it has been submitted, accepted and reviewed for a journal

Maybe Arts & Humanities don't need the wider public to have immediate access to our materials – the need for timeliness is one of the main reasons pre-prints started in the sciences – lag in time between peer review and publishing.

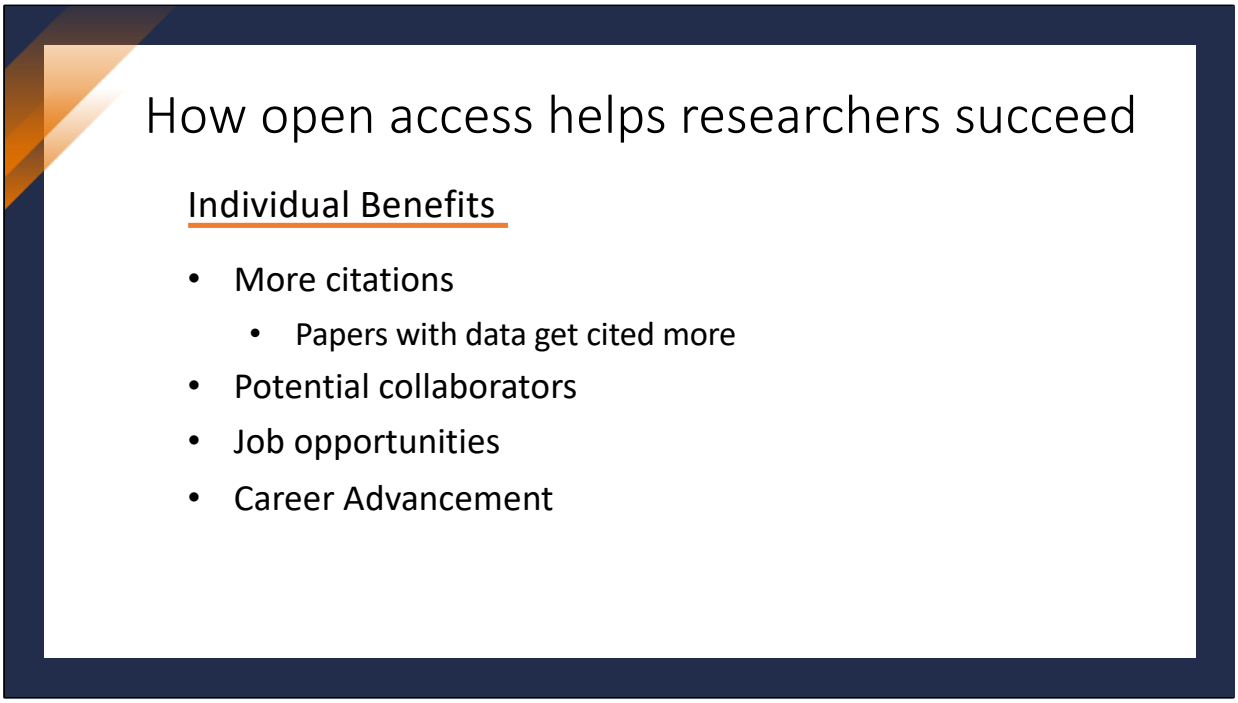
BUT making your work OA (especially early and often) might mean **more folks can access and then identify your work with your name, and you have a clear digital trail** showing how long you've been developing your work.

sharing our ideas and encouraging feedback and collaboration early on in the research process makes our research stronger.

Example of how pre-prints could work in AH: repository of preprints in the humanities – making available in a very quick timescale new research ideas that are not yet developed into their full, final form and encouraging scholars to offer constructive feedback and dialogue to aid the research process, out of which extended articles can then be developed.

Or think of sharing all supporting materials such as presentations, posters, blogs, data papers etc. and consider using social media for wider outreach,

I also want to point out that pre-prints are good for disseminating for equity - giving more access to the content than behind paywall



Here are some Potential benefits for being open in your research. The benefits certainly are competing with risks and pressures - all structural (Institutions for prestige, journals for tying researchers up with impact factors - journal prestige and competition)

Take a look at these benefits… how could they apply to your discipline.

Being Open helps…..

- Potential collaborators (more readers)
- Career advancement - with your work open you will have an opportunity to find new projects and new collaborators

open access publishing might make your work available to a broader readership simple by virtue of availability, which could be useful for them as they look at a variety of careers.

I want to point out that openness is not 'all-or-nothing'. Not all researchers are comfortable with the same level of sharing, and there are a variety of ways to be open (the paper cited in the slide notes has a great list of open-ways). Openness can be thus defined by a continuum of practices, starting perhaps at the most basic level with openly self-archiving pre-prints and post-prints. Fully open research is a long-term goal to strive towards.

==========

Papers with Data get cited more: point out a couple of studies in the slide notes:

Piwowar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing detailed research data is associated with increased citation rate. PloS one, 2(3), e308. https://doi.org/10.1371/journal.pone.0000308
- using cancer microarray clinical trial publications - Publicly available data was significantly associated with a 69% increase in citations,

Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. https://doi.org/10.1371/journal.pone.0230416
- journal articles published by PLOS and BMC, positive and significant relative gain in citation counts in general for a paper with data available in a repository.

Tools that help being more Open

- Pre-Print servers
  - https://en.wikipedia.org/wiki/List_of_academic_publishers_by_preprint_policy
- Sherpa/Romeo
  - https://v2.sherpa.ac.uk/romeo/
- DOAJ: Directory of Open Access Journals
  - https://doaj.org/

Publish where you want (with the prestige publishers) and archive openly in a pre-print repository
Again, Pre-print is a version of your article before it has been submitted, accepted and reviewed for a journal

You need to be aware of publisher policies before you make your pre-print available…… The first link here on the slide is a list of publisher policies

Sherpa/Romeo is another site that breaks down publisher (journal) policies in regards to pre-prints and post-prints
Publishers' policies and conditions regarding open access archiving (also known as self-archiving) and other rights and permissions options can be confusing. Sherpa RoMEO, a directory of publishers' copyright policies, lays out the details clearly.
OR deposit a post-print in an institutional repository (authors' version after peer review, but before journal editing).

If you are interested in publishing in Open journals, you can check out the Directory of Open Access Journals
The DOAJ
database contains over 16 500 peer-reviewed open access journals covering all areas of science, technology, medicine, social sciences, arts and humanities.
DOAJ's mission is to increase the visibility, accessibility, reputation, usage and impact of quality, peer-

reviewed, open access scholarly research journals globally, regardless of discipline, geography or language.

Tools that help being more Open

- Quality checks on open publishing
  - http://whyopenresearch.org/journals
- Your Library!!!
  - Institutional Repository
    - UVA – Libra (Open works and Open Data)
  - Open Journal Management System
    - UVA – Aperio (open press supported by UVa Library)

This site “whyopenresearch.org” Walks you through resources to find a suitable open access journal and then a series of checks to make sure it is good for your research.

Many Library have institutional repositories: This would not only provide open access to those articles, but also give the authors freedom to publish in any journal of their choice, open access or not.
Or disciplinary repositories

Using an institutional repository
The scholarly outputs of researchers are immediately visible, accessible and discoverable online. This means that an increased number of readers can convert into an increased number of citations. Adding research to an open access “institutional” Repository preserves that institution’s publishing output. And if a persistent identifier is added to all material posted (DOI), means that the URL never changes and is resolvable even if the content is migrated to a new system.

For data to be managed (for "long-term") and made accessible in a sustained way, it should be deposited in a "trusted" repository".
You want to choose a FAIR repository
Depositors and re-users should be assured that data are authentic, retrievable, and documented enough to understand the context of how they were created - and have a license that clarifies the conditions of reuse.

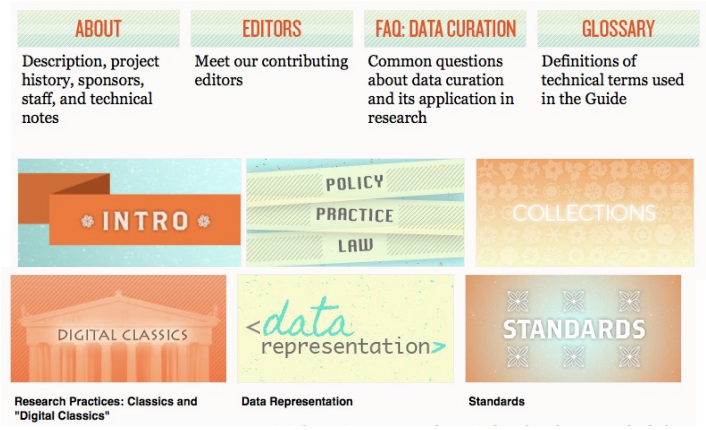
Use open-source journal management software to reduce the costs about as far as possible.
Move paywalled journals to open journals

Create new journals

UVa has Aperio, check w/ your institution (Library or Press)

Digital Curation Guide

A community resource guide to data curation in the digital humanities

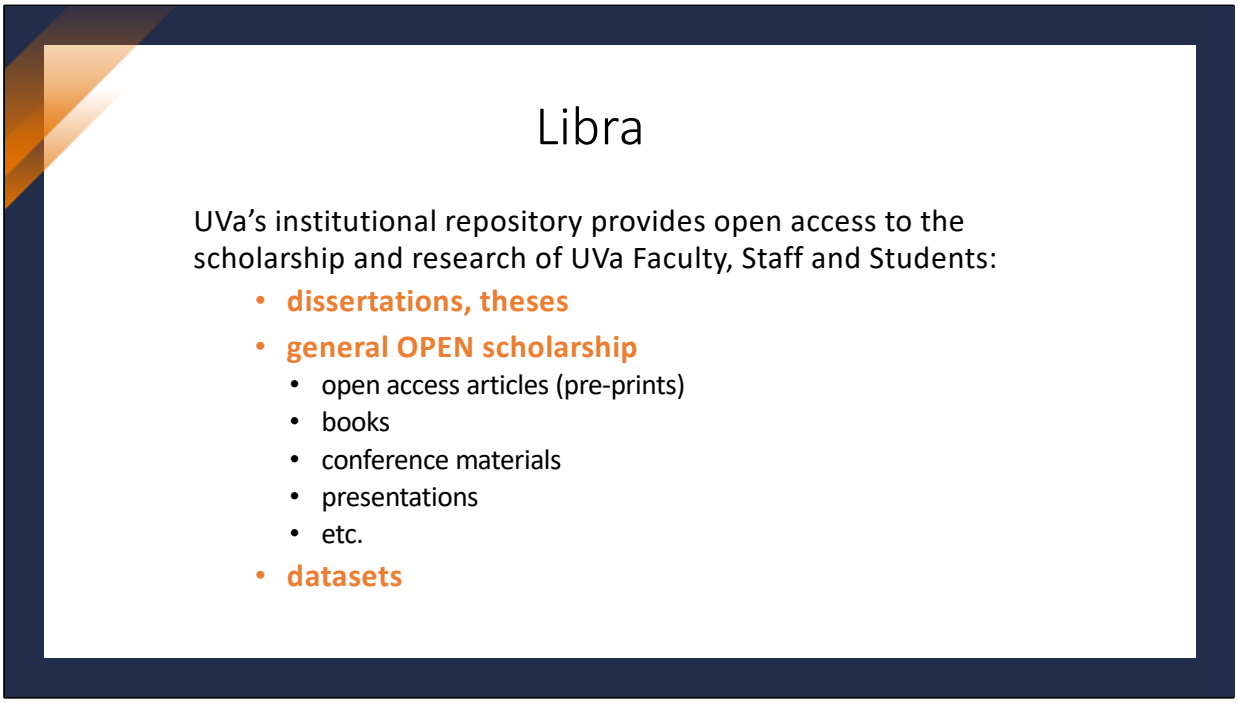


http://guide.dhcuration.org/

If you want more information on organizing and sharing digital humanities data – an older, but still useful guide.

The goal of the DH Curation Guide is to direct readers to trusted resources with enough context from expert editors and the other members of the research community to indicate to how these resources might help them with their own data curation challenges.

understanding the essentials of data curation for the humanities. The expert-written introductions to key topics include links to important standards, documentation, articles, and projects in the field, annotated with enough context from expert editors and the research community to indicate to newcomers how these resources might help them with data curation challenges.



What is Libra?

Libra is UVa's institutional repository for scholarly works from UVa's faculty, staff and students. It provides open access to publicly available content such as papers, conference proceedings, presentations, posters, reports, pre-prints, open version of scholarly articles, Electronic Theses and dissertations and datasets.

Benefits of Submitting Works to Libra

By submitting works to Libra, UVa's scholarship is discoverable and downloadable by other scholars worldwide. Other benefits include:

- Immediate availability of UVa scholarship, without the delays necessary for binding and cataloging paper copies. "Free" from subscriptions
- Persistent URL via Digital Object Identifier (DOI) registration provides a simple, durable link to include in your CV, on your personal web site, and other places where you share your scholarship.
- Optional ORCID iD linking to connect your scholarship to you via persistent digital identifier, in addition to your name.
- Preservation of works in Libra as part of the record of UVA scholarship.

Where is it?

Since Libra is really three separate servers, we have created one landing page that gets you to all the information for the repositories; access to deposit AND searching.

The Libra landing page has links to more information on the repositories, including links to the various repositories

Who can use it?

Any faculty/staff or student can deposit works in LibraOpen or LibraData

LibraETD is for degree granting departments (degrees) as well as optional Honor's programs and Distinguished Major's

LibraOpen is for anyone at UVa to share their work.: papers behind paywall?? – deposit an open version in Libra. Conference online? – submit your poster or recorded presentation or paper

LibraData is UVa's institutional version of Datavese, created by Harvard. Is for data and other products of research that supplement a paper OR for research that needs to be shared per funder requirements.

Each repository has a "Checklist".

How to search?

Since we have three separate systems, all searching is done in Virgo, UVa Library catalog.

[Although you can search datasets directly in UVa's Dataverse.]

To search only items in Libra (all three of the systems), use the search box on the Libra Landing page: http://libra.virginia.edu

Discussion?

Sherry Lake
Scholarly Repository Librarian
libra@virginia.edu

Slides and notes available: https://doi.org/10.18130/c25f-nc93

References

ALLEA. *Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group E-Humanities.* Edited by Natalie Harrower, Maciej Maryl, Beat Immenhauser, and Timea Biro. 2020, https://repository.dri.ie/catalog/tq582c863. [Note: Previous link not resolvable - Copy of report here https://virginia.box.com/s/ej6v5in5018rssbmepwy3ujvfbk9ho5i]

Edwards, Caroline. "How Can Existing Open Access Models Work for Humanities and Social Science Research?". *Insights* 27 (1): 17–24. 2014, http://doi.org/10.1629/2048-7754.135

Grant, Rebecca. *Opening up Humanities Scholarship: Open Humanities Data on the Routledge Open Research Publishing Platform*. https://oaspa.org/wp-content/uploads/2022/09/ROR-OASPA-2022-Poster-FINAL-Rebecca-Grant.pdf. OASPA Online Conference on Open Access Scholarly Publishing 2022.

Nowviskie, Bethany. “Why, Oh Why, CC-BY?” *Bethany Nowviskie Blog*, 11 May 2011, http://nowviskie.org/2011/why-oh-why-cc-by/.

Posner, Miriam. “Humanities Data: A Necessary Contradiction.” *Miriam Posner's Blog*, 25 June 2015, https://miriamposner.com/blog/humanities-data-a-necessary-contradiction/.

Rutledge Open Research. *Open Data and Accessible Source Materials Guidelines*, 2022, https://routledgeopenresearch.org/for-authors/data-guidelines.

Suber, Peter. “Why Is Open Access Moving So Slowly In The Humanities?” *Blog of the APA*, 8 June 2017, https://blog.apaonline.org/2017/06/08/open-access-in-the-humanities-part-2/.

Wilkinson, Mark D., et al. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data*, 15 Mar. 2016, https://doi.org/10.1038/sdata.2016.18.