Library Scholarship   James E. Brooks Library

Fall 12-2020

ScholarWorks Procedures and Policies

Maura Valentino
Central Washington University, maura.valentino@cwu.edu

ScholarWorks Open Institutional Repository

1. Library Mission:

The libraries of the Central Washington University, in keeping with the university's mission, foster learning and academic excellence through diverse collections, services, and teaching. This is accomplished by creating a culture of inquiry, research, and lifelong learning in the academic and regional communities. The libraries support and extend the Central Washington University's student-focused and global mission.

2. About ScholarWorks

The Central Washington University ScholarWorks is a central repository for scholarship and creative work contributed by the students, faculty and staff of the university with the goal of preserving the history and scholarship of the university and expanding the visibility of the university to the regional and global community. University community members are encouraged to contribute works such as articles and books, journals, presentations, conference proceedings, creative activities, media, theses, and archival records of the university. ScholarWorks is a university-wide initiative managed by the University Libraries.

ScholarWorks Organization

ScholarWorks has 6 overarching collections:
- Faculty Scholarship and Creative works
  Contains articles, white papers and other works by CWU faculty. This collection has collections contained within it, divided by College.
- Student Scholarship and Creative works
  Contains Theses, Capstone projects, Student projects and PULSE, a student publication
- Community Projects
  Contains collaborative collections between CWU Libraries and community organizations such as the Ellensburg Public Library.
- University Archives and Special Collections
  Contains digital versions of some of the University Archives and Special Collections.
- Journals
  Contains the Journal of Math Circles published by the Math Department
- Conferences and Events
  Contains the SOURCE collection and the Brooks Library Poetry Series

3. Collection Development Policy:

o **Who can contribute:** ScholarWorks includes work created by individuals or groups affiliated with a CWU department, program, research team, working group, collaborative project, or conference, such as:
  ▪ CWU faculty, researchers, and staff

- CWU departments, schools, and colleges
- Undergraduate and graduate students with faculty sponsorship
- Other University-affiliated entities will be considered on a case-by-case basis.

- **What kinds of materials:**

  ScholarWorks preserves and makes publicly available scholarly, scientific, professional, and cultural works created within the CWU community including:

  - Dissertations and theses
  - Capstone Projects
  - Faculty and student scholarship (articles, presentations, multimedia)
  - Documentation of CWU-hosted conferences and events (papers, presentations, programs, videos, proceedings)
  - Grey literature (technical reports, research reports, working papers)
  - Digitized special collections
  - Digitized archival collections
  - Open access journals
  - Other types of materials will be considered on a case-by-case basis, including:
  - Books and book chapters
  - CWU archival records (policy documents, meeting notes, reports)
  - OER course materials
  - Data sets and/or descriptive metadata
  - The CWU Libraries reserve the right to deny or remove materials not within the scope of the collection development policy.

- **Priorities:**
  - Theses and dissertations
  - Capstone Projects
  - Journals and conferences
  - Honors theses and capstone projects
  - Conference, symposia, and other event content (faculty and student)
  - Faculty scholarship, including gray literature, articles, pre-prints/post-prints, book chapters, conference papers, presentations
  - Institutional documents and publications
  - Photographs, art, creative collections (faculty and student)
  - Performances, lectures (faculty and student)
  - Student media publications
  - Peer-reviewed journals
  - Data sets
  - Resources unique to the Central Washington University

Size and Format:

- ScholarWorks imposes no maximum file size limit. However, most Internet connections will not support an upload of more than two gigabytes in size. For large files, or for collections totaling more than two gigabytes in size, please consult with the Scholarly

Communications Librarian to determine the best ingest method. For extremely large datasets, you may instead be refered to a subject-specific repository.

- File Formats
- *Supported:* When an item's format is public and open, as is the case with formats such as Adobe PDF, HTML, or JPEG it is categorized as a "supported format." Items in this category can be used in the future through migration or emulation and the MSUL makes a commitment to do so. Examples:
  - Adobe PDF (pdf)
  - Compressed files (.zip, .tar, .rar)
  - GIF (gif)
  - HTML (html, htm)
  - JPEG (jpg, jpeg)
  - MPEG (mpa, abs, mpeg)
  - PNG (png)
  - Rich Text (rtf)
  - SQL
  - Text (txt)
  - WAV (wav)
  - XML (xml)

The ScholarWorks team hopes to provide support for as many file formats as possible. Over time, items stored in ScholarWorks will be preserved as is, using a combination of time-honored techniques for data management and best practices for digital preservation. As for specific formats, however, the proprietary nature of many file types makes it impossible to make guarantees. Put simply, our policy for file formats is this:
- Everything put in ScholarWorks will be retrievable.
- We will recognize as many files' formats as possible.
- We will support as many known file formats as possible.

We will preserve files, and in cases where formats are native to tools supported by CWU Information Systems, we will provide you with guidance on converting your files into formats we do support. It is also likely that for extremely popular but proprietary formats (such as Microsoft .doc, .xls, and .ppt), we will be able to help make files in those formats more useful in the future simply because their prevalence makes it likely tools will be available. Even so, we cannot guarantee this level of service without also having more information about the formats, so we will still list these formats as "known", not "supported."

Data Sets

The data must be produced, submitted or sponsored by CWU faculty or researchers.
- The data should be complete and ready for distribution.
- Classified, confidential, and/or restricted data are not accepted.
- The author/owner must be willing and able to grant CWU the right to preserve and distribute the data.
- Zipped and/or tarred files are discouraged, but may be used in the event that a dataset is too large or contains many individual files that should be distributed as a

bundle. Please contact your Scholarly Communications Librarian if you plan to submit zipped and/or tarred files.

• The dataset may contain multiple files. Individual data files (including zipped/tarred files) may not exceed 2GB.

• Each submitted dataset should contain:

a) A 'readme' file listing:

- The contents of the submitted dataset (i.e., file names, formats and sizes
- of each file).
- A list of software used to produce, render and compress the data (where
- applicable).
- Where in the research process the data falls (e.g. raw/unanalyzed data,
- processed/analyzed data, rendered/visualized data).

b) Submitters are responsible for adequate description of the contributed dataset for the purposes of discovery and other informational purposes (e.g., funder names, grant numbers, how the data was collected).

c) It is recommended that datasets be documented sufficiently so that other knowledgeable researchers can find, understand and use the data.

Metadata

ScholarWorks will use a hybrid form of the Dublin Core metadata standard, which is the same standard used to describe CWU Digital Collections objects. The depositor of a digital work is encouraged to provide a basic level of metadata during the deposit workflow, such as author(s), abstract, subject keywords, citation information, ORCID (Open Researcher and Contributor ID), and DOI (Digital Object Identifier). More detailed metadata may be entered during the IR approval process to provide more robust cataloging information that aids in searching and indexing.

4. Copyright/permissions policy:

The Central Washington University claims no ownership over the content but maintains a non-exclusive right to make the materials accessible for non-commercial use. The non-exclusive right does not interfere with the copyright holders' rights or ability to reuse content. Copyright is held by the original authors/creators or the publisher if rights were transferred. The Scholarly Communications Librarian may assist with specific questions about copyright and permissions, but the primary responsibility of conducting copyright clearance falls to the contributor. Items in the repository hold a Creative Commons License.

5. Digital Preservation Policy:

The Libraries are committed to preserving and providing access to the scholarly content in ScholarWorks. Some file formats may cease to be supported as a result of changes in technology beyond the control of the Libraries. Bepress, the publisher of the Digital Commons platform, employs multiple preservation strategies to ensure the long-term integrity of deposited materials. These measures include: provision of stable URLs, robust infrastructure, daily and weekly offsite

backups, dedicated staff, scheduled maintenance, software testing and development, file format migration and emulation, and OAIS and LOCKSS compliance.

Digital Initiatives and Special Collections and Archives will promote the use of standard and open content formats that meet community-accepted digital preservation standards, are widely supported on a number of platforms, and that can be preserved and migrated forward to new preservation formats over time. The preferred file type for documents housed in ScholarWorks is Adobe Acrobat PDF. When materials are submitted in Microsoft Word (doc and docx) or Rich Text Format (rtf), ScholarWorks will automatically convert them to searchable PDF, retaining copies of the original ScholarWorks for preservation purposes. Other file types will not be automatically converted to PDF as part of the submission process. PDFs that are uploaded (whether searchable or not) will not be altered in the submission process. For documents not in PDF format, ScholarWorks staff recommends that a PDF version be submitted along with the original whenever possible. The PDF should have all pages oriented in the same direction, be the same size and be as straight as possible. PDFs should have the ability to be put into a book reader.

6. Access and Use Policy:

Users are encouraged to cite and link to digital content in ScholarWorks and are free to do so without asking for permission. Depending on the source of the digitized work, licenses or other contractual terms may restrict further distribution or other uses. ScholarWorks does not have authority to grant or deny permissions, as copyright remains with the authors. If permission is required, Special Collections and Archives should be consulted. When any permissible use is made of the materials, we request that ScholarWorks be attributed as the source of the digital contents and that links to the item on ScholarWorks be included where appropriate. If you have any questions about access and use, please do not hesitate to contact Special Collections and Archives.

7. Privacy Policy:

ScholarWorks respects the privacy of all contributors, visitors and users of its services. Information about privacy in different areas of ScholarWorks is given below.

- o **Assessments:**

  Information and data obtained by ScholarWorks in support of assessment of services, collections, resources, etc., or in support of research related to repository services, are considered confidential and will not be shared except in aggregates or with the express permission of participants, to protect privacy.

- o **BePress Analytics**

How can you know the difference your scholarship is making without knowing where it's being read? We offer detailed, exportable analytics that show how readership changes over time,

whether for a department, an author, or the entire institution. Our industry-leading readership tools show the specific organizations and regions that are engaging with your work (and how they're finding it), so you can better identify research trends, potential funders, and collaborations.

8. Submission and Hosting Guidelines:

- **Theses/Dissertations/Projects:**

  Students are required to submit their own thesis to the correct community. The work must be submitted electronically. With the submission, the contributor must also sign and submit a permissions form, and the committee's approval/signature page, in which signatures will be redacted, before the work will be uploaded to ScholarWorks.

- **Individual Articles/Publications:**

  Faculty may upload published articles to the repository if the articles are Open Access or if the journal allows a copy of the article to be uploaded into the repository. The Scholarly Communications Librarian will track faculty publications and upload publications as copyright allows. Faculty may also give a list of publications to the Scholarly Communications Librarian for possible inclusion into ScholarWorks.

- **Journals:**

  Those wishing to publish open access journals through ScholarWorks should contact the Scholarly Communications Librarian. Any academic unit is welcome to start a journal and have it hosted in ScholarWorks. An example of a CWU journal may be found at: https://digitalcommons.cwu.edu/mathcirclesjournal/

- **Conferences and Symposia:**

  Conference and Symposia that are affiliated with CWU may have their materials uploaded to ScholarWorks. Contact the Scholarly Communications Librarian.

9. Withdrawal and retention guidelines:

ScholarWorks is meant to be a permanent scholarly record. Once an item is deposited in the repository a persistent URL will be generated and a citation to the item will always remain. Removing content is discouraged. However, authors may request that the community administrator remove their files. The community administrator must approve all such withdrawals in consultation with the author. Reasons for removal may include copyright infringement, plagiarism, or falsification of data. No files will be removed by the community administrator without the author being informed.

Students may elect to remove their work after a time period of 1, 3 or 5 years.

Updated versions of content may be posted, and can be linked to the first posted version. Earlier versions may be removed, but we encourage authors to maintain these versions in the repository, as a record of the development of any posted contents.

10. Contact information:

Please contact the Scholarly Communications Librarian with any questions or concerns.

11. Digital Library Policies

CWU Digital Libraries provides a wide variety of digital content from multiple partners and contributors. From mass digitization projects to targeted subject collections of digitized and born-digital materials, the Digital Libraries Division offers state-of-the-art infrastructure, management, storage, access, and preservation services for digital knowledge. In support of the CWU Libraries Mission, the CWU Libraries' Digital Collections is dedicated to the long-term collection, production, maintenance, delivery, and preservation of a wide range of high-quality digital resources and services for the CWU Community and users throughout the world. The goal of this policy is to create a consistent, structured approach to evaluating digitization projects so that collections selected for digitization will have the greatest impact on research, scholarship, and learning at CWU, and will best support and reflect the mission and goals of the Library and the University.

Materials to be considered under this policy include text and manuscript documents, photographs, fine art, illustrations, sound recordings, video recordings, 3-D objects, and other types of materials. Collections considered under this policy may come from the Library's general collections, CWU Special Collections and Archives and community partners.

CWU Libraries' Digital Collections exist to:

- Acquire, preserve, and provide access to digital resources required by the CWU community to successfully execute the mission of CWU as a major public research university deeply committed to advancing educational excellence and preparing students to become thoughtful, engaged citizens of the world
- Provide an optimum research environment to faculty, staff, students, and community members by enabling open access to significant research collections
- Support open access to and long-term preservation of the scholarly output of CWU to the best of its ability
- Encourage long-term preservation of content contributed by partners, researchers, and institutions
- Provide digital preservation support for non-CWU community organizations in the nine counties surrounding CWU, who otherwise do not have access to preservation support

The CWU Libraries' Digital Collections have a broad audience interested in utilizing the digital resources, including:

- The CWU Extended Community (comprising students, faculty, staff, alumni, and administrators)
- Contributing partners and their constituents
- The larger academic community and researchers
- The general public throughout the world

Each item must:

- Be in the public domain, or have documented, non-revocable permission granted by the copyright holder, or be allowed by Section 108 (17 U.S.C. § 108)
- Be intended for public viewing and use
- Be complete, such as an entire publication, article, etc., and not a "part" such as an abstract, forward, or title page
- Meet the standards required for long-term digital curation
- Be in a standard format accessible through current file viewers or have a documented conversion path to move the format into a standard format
- Be intended for permanent storage in the CWU Libraries' Digital Collections
- Be in current use and have potential future use
  - The collection has the potential for or is already high-use.
  - There is a defined set of users or anticipated users for the collection.
  - The collection is rare or difficult to access.
  - The collection will attract use from both Georgetown and external communities.
- Have a relationship to other collections
  - The materials form a coherent collection, fill gaps in existing collections, or complement existing collection strengths.
  - There is potential for collaborative digital collection building or other collaborative relationships with other institutions.
- Have Prestige, Innovation, Inspiration, and Funding
  - The collection brings institutional prestige, distinctiveness, and/or recognition to the Library or University.
  - The collection provides an opportunity for technological innovation.
  - The collection provides inspiration for and ability to engage in transformative teaching, research, and/or learning.
  - Digitizing the collection offers the potential to attract funding.

Criteria for Digital Projects: Technical

In addition to assessing proposed digitization projects from a collection development perspective, the collection will be reviewed to determine how difficult it will be to digitize the collection with respect to: preservation, metadata, technology, and rights/privacy. If it is not feasible to digitize a collection due to technical concerns, the project will be turned down or deferred.

1. Preservation

Only collections where the gains from the resulting digital collection clearly outweigh the potential harm to the original materials during digitization will be considered. The following questions provide a structure for evaluating a collection from a preservation perspective:

- o Are the materials at risk of being lost due to poor condition or obsolete media format?
- o Will reformatting put the original materials at an unacceptable risk of damage?
- o Is conservation work needed prior to digitization?

2. *Metadata*

Ideally, collections will be organized and described before they are digitized, though it is likely that additional metadata will need to be created for a digital collection. While different metadata schema will be used to support different digital collections, the metadata for all projects must adhere to community agreed-upon best practices. The following questions provide a structure for evaluating a collection from a metadata perspective:

- o Have the materials been organized and processed?
- o What metadata is already available?
- o What level of metadata is needed for the content of the collection to enhance the discovery of, access to, and use of the digitized materials?
- o What metadata work is needed to support the project?

3. *Digitization*

Technical expertise and specialized equipment are required to reformat a physical collection to a digital collection. While different technical standards may be used to support different digital collections, the standards for all projects must adhere to community agreed-upon best practices. The following questions provide a structure for evaluating a collection from a digitization perspective:

- o It is technologically possible to digitize the items with resources available in the Library?
- o Can the digital objects be reformatted and stored according to relevant technical specifications, guidelines, and best practices?
- o Are the derivative asset file formats supported by the Library's digital collections platform?
- o Will the Library's digital collections platform allow the user to use the collection effectively?

4. *Rights & Privacy*

The Library has a strong preference to make its digital collections freely available to a worldwide audience for research and educational purposes. The following questions provide a structure for evaluating a collection from a rights and privacy perspective:

Copyright

- o Are the materials clearly in the public domain?
- o Does the Library have permission to digitize the materials? If not, can the rights holder(s) be located to secure permission?
- o Are there other options that would allow digitizing the materials and making them freely available?

Privacy

- o Does the collection contain materials with personal information or other content which, if made public, would violate privacy laws or University policies?
- o Is there any other reason why it would be unwise to make the collection, or certain materials within the collection, public?

Digitization and Scanning standards

Text (Books)					
Image Types	Bit Depth	Color Space	Resolution (	Scale	File Format
B&W Text Only	1-bit	Bitonal	600ppi	100% (1:1)	Tiff (uncompressed or loss
B&W Text with Illustrations	8-bit	Grayscale	400ppi	100% (1:1)	Tiff (uncompressed)
Text w/ Continuous-Tone Im	24-bit	RGB	400ppi	100% (1:1)	Tiff (uncompressed)

Documents (Handwritten, Ledgers, Receipts)					
Image Types	Bit Depth	Color Space	Resolution (	Scale	File Format
B&W Text Only*	1-bit	Bitonal	600ppi	100% (1:1)	Tiff (uncompressed or loss
Handwritten Documents, Led	8-bit	Grayscale	400ppi	100% (1:1)	Tiff (uncompressed)
Correspondence, Letters	24-bit	RGB	400ppi	100% (1:1)	Tiff (uncompressed)

* We recommend scanning handwritten documents in grayscale or color.

Photographs (Prints)					
Image Types	Bit Depth	Color Space	Resolution (	Scale	File Format
B&W Photographs	8-bit	Grayscale	600ppi	100% (1:1)	Tiff (uncompressed)
Color Photographs	24-bit	RGB	600ppi	100% (1:1)	Tiff (uncompressed)

Photographs (Negatives and Slides)					
Image Types	Bit Depth	Color Space	Resolution (	Scale	File Format
B&W Negatives & Slides	8-bit	Grayscale	5000 pixels a	100% (1:1)	Tiff (uncompressed)
Color Negatives & Slides	24-bit	RGB	5000 pixels a	100% (1:1)	Tiff (uncompressed)

Large Format (Posters and Maps) Above 17 x 24 inches (A2)					
Image Types	Bit Depth	Color Space	Resolution (	Scale	File Format
B&W Maps/Posters	8-bit	Grayscale	5000 - 10000	100% (1:1)	Tiff (uncompressed)
Color Maps/Posters	24-bit	RGB	5000 - 10000	100% (1:1)	Tiff (uncompressed)

Newspapers					
Newspapers should be scanned and digitally preserved in accordance with the national standards set by the					
Image Types	**Bit Depth**	**Color Space**	**Resolution (**	**Scale**	**File Format**
Microfilm	8-bit	Grayscale	400*	100% (1:1)	Tiff 6.0 (uncompressed)
Physical Pages	24-bit	RGB	400	100% (1:1)	Tiff 6.0 (uncompressed)
Monitor equipment performance by capturing a standards-based target film strip or color image at the start					
Split dual images into individual newspaper images as necessary.					
Deskew images with more than 3% skew.					
Crop page image files to the edge of the newspaper, retaining the original edge and up to a quarter inch bey					
Produce images that have exactly the same dimensions, spatial resolution, skew, and cropping as the image					
*For microfilm, scan at 8-bit grayscale with a resolution of 400 dpi, if possible; otherwise between 300 and					
Moving Image (Film and Video)					
Original Format	**File Format**	**Codec**	**Audio**	**Resolution**	
16 mm film	QuickTime (.	Uncompress	48 kHZ, 16-b	Standard Definition (720x480)	
VHS, SVHS, Betamax, Betaca	QuickTime (.	DV50	48 kHZ, 16-b	Standard Definition (720x480)	
Audio					
Original Format	**Bit Depth**	**Resolution**	**File Format**		
			Waveform Audio File Format (WAVE)		
Cassette	16-bit	44.1 kHz	.wav		
			Waveform Audio File Format (WAVE)		
Reel-to-Reel, Analog Disk	24-bit	48 kHz to 96	.wav		