

Digital Collection Development Policy

Building and Sharing Nationally Recognized Digital Collections

The University of Houston Libraries' Digital Collections are central to the Libraries' vision of becoming a nationally recognized leader among research libraries. In alignment with the University of Houston's the University of Houston Libraries' Strategic Plans, Digital Collections support the research and teaching needs of UH faculty and students, fulfill the needs of campus and community constituencies, collect, present, and preserve digital files of material appropriate for scholarly research, and promote open access to materials of scholarly and historical significance. UH Libraries is committed to ensuring that digital collections reflect the diversity of the UH community and will routinely assess for content inclusive to a multicultural campus and city.

UH Libraries Digital Repositories

This policy highlights selection criteria for multiple digital repositories managed by UH Libraries in collaboration with outside stakeholders, including the Texas Digital Library and HathiTrust.

Repository	Description
UH Digital Collections	These materials document the University of Houston, city of Houston, and state of Texas, as well as other historically and culturally significant topics. The repository is managed by UH Libraries.
Cougar Research Open Access Repositories (ROAR)(Hosted by TDL)	<p>The UH Libraries' Institutional Repository (IR) accepts content reflecting the scholarly and administrative output of the university. It makes available digitized and born-digital scholarship and creative works from faculty, students, and staff. The IR also contains digitized and born-digital University publications. The IR runs on the DSpace platform and is maintained in partnership with the Texas Digital Library.</p> <p>The UH Dataverse stores and makes available research data from UH faculty, students, and staff. It operates on the Dataverse platform and is maintained in partnership with the Texas Digital Library.</p>
Audio/Video Repository	The Audio/Video Repository accepts AV format collections owned, managed, or generated by UH Libraries and the University of Houston community. The repository is managed by UH Libraries.

	<p>Collections that may otherwise have been added to the UH Libraries' IR will be accepted in the AV Repository if the materials require streaming media for better user experience, or if download restrictions or fine-grained access controls are needed due to the nature of the AV materials, their copyright status, etc.</p>
Third-party Repositories	<p>UH Libraries may at times contribute digitized content to third-party repositories managed outside UH, including open access and subscription-based systems. Use of open access repositories is encouraged to best support the educational and research mission of the Libraries. Contributing to these repositories may currently remain aspirational.</p> <p>The HathiTrust Digital Library is a large-scale digital preservation repository focused on preserving and providing access to book and journal content digitized from member collections through a number of means, including digitization by Google, the Internet Archive, and through local initiatives. The repository is maintained by HathiTrust.</p> <p>The Portal to Texas History collects digitized “rare, historical, and primary source materials from or about Texas” from over 400 content providers. The Portal is created and managed by the University of North Texas Libraries.</p> <p>The Digital Public Library of America works with a national network of content providers, including the Portal to Texas History, to make digital materials aggregated from libraries, archives, museums, and other cultural institutions across the country available to all in a single repository.</p>

Content Deposit Locations

Content is deposited into one of the existing digital repositories based on its content type. Note that materials may also live in multiple repositories, based on the functions of those particular repositories. As services and functionality around repositories expand, the defined content areas for any one repository may change or merge over time.

Type of content	Definition	Repositor(ies)
UH Libraries collections/materials	<ul style="list-style-type: none"> Digitized versions of analog and born-digital library collections, including print monographs and serials, archival materials, maps, photographs, and 	<p>Primary:</p> <ul style="list-style-type: none"> UHDC <p>Other:</p>

	musical scores, held by the UH Libraries	<ul style="list-style-type: none"> • Third-party Repositories
UH faculty, staff, and student scholarship and creative work	<ul style="list-style-type: none"> • Electronic theses, dissertations, and graduate capstone projects, completed by UH graduate students in fulfillment of degree requirements • Previously published, peer-reviewed, or juried works—e.g., books, journal articles, music scores, poetry, fiction and creative non-fiction, art, music recordings, and dance—with permission of rights-holder(s) • Prepublication pre-/post-prints with permission of rights-holder(s) • Grey literature (e.g., technical reports, white papers, research posters, or symposia proceedings) and other unpublished scholarly works • “Unpublished” creative work • Learning objects, including instructional materials and primary sources organized for teaching purposes 	<p>Primary:</p> <ul style="list-style-type: none"> • Cougar ROAR
Digitized and born-digital University Archives material	<ul style="list-style-type: none"> • Electronic administrative records produced by university units to document actions and/or decisions • Student organization digital objects, including rosters, meeting minutes, and event publicity materials • Web harvesting of website material of enduring value 	<p>Primary:</p> <ul style="list-style-type: none"> • Cougar ROAR
Audiovisual material	<ul style="list-style-type: none"> • Digitized versions of analog or born-digital audio and video collections, held by the UH Libraries or generated by UH Libraries employees/UH Faculty 	<p>Primary:</p> <ul style="list-style-type: none"> • Audio/Video Repository <p>Other:</p> <ul style="list-style-type: none"> • Cougar ROAR
Datasets	<ul style="list-style-type: none"> • Raw or processed collections of data, usually presented in a structured form, that are the product of one phase of 	<p>Primary:</p> <ul style="list-style-type: none"> • Cougar ROAR

	<p>research and form the basis for further work</p> <ul style="list-style-type: none"> ○ Note: Dataset storage will comply with grant or project guidelines; storage and service provided after the life of the project should be re-evaluated in alignment with library resources. 	
--	---	--

UH Digital Collections Selection Guidelines

Collections and items in the UH Digital Collections should, first and foremost, generate national recognition for the University of Houston and the UH Libraries.

To be included in the Digital Collections, collections and items should also:

- Be of significant research and/or teaching interest
- Align with University of Houston strategic priority areas, especially Energy, Health, or Arts
- Exist nowhere else as digital content that is easily accessible and/or of comparable quality
- Meet existing or anticipated demand

Repository	Criteria for Inclusion
UH Digital Collections	<ul style="list-style-type: none"> ● Collections that are worthy of national recognition. They address critical elements including: <ul style="list-style-type: none"> ○ Supporting research and/or teaching at the University of Houston ○ Aligning with one or more of the University of Houston's strategic priority areas ○ Amplifying traditionally underrepresented voices in the existing collection ○ Aligning with popular research trends or issues ○ Supporting open access initiatives ○ Making collections of unique or rare items of historical or cultural significance more accessible ○ Complementing or completing existing digital collections (local or external) ○ Complementing collections selected for commercial digital publication

	<ul style="list-style-type: none"> ● Collections that fulfill important local needs. They address critical elements including, for instance, important local projects and university connections: <ul style="list-style-type: none"> ○ Supporting a program or project at UH Libraries or the University of Houston (e.g., a conference) ○ Has potential for external funding <p><i>Technical criteria (for materials needing digitization):</i></p> <ul style="list-style-type: none"> ● Can be digitized without damaging original ● Has a desirable preservation outcome (e.g., reduced handling of fragile original) ● Will satisfy requests for digital surrogates ● Would provide access to otherwise inaccessible content ● Providing online access will not violate existing copyright or intellectual property laws
Cougar ROAR (Hosted by TDL)	<p>The UH IR provides access to content created by a University of Houston-affiliated individual or unit, including (but not limited to):</p> <ul style="list-style-type: none"> ● Faculty, staff, and student scholarship and creative works ● College and departmental reports and publications ● University publications <p>The UH Dataverse is a platform for publishing and archiving datasets (and other data products) created by faculty, staff, and students at the University of Houston. Users also have the ability to restrict access to data as desired.</p> <ul style="list-style-type: none"> ● The UH Dataverse is configured to accept any particular type or subject of research and all file formats are supported regardless of the research type. ● Potential types of research data include: <ul style="list-style-type: none"> ○ scientific experiments (to include social sciences and humanities) ○ input data and simulations results ○ derived data (from processing or combining “raw” or other data) ○ canonical or reference data (gene sequences, chemical structures, etc.) ○ accompanying material, observations, or ephemera ○ oral history recordings

	<ul style="list-style-type: none"> ● The UH Dataverse does NOT accept content that contains confidential or sensitive information, and requires that contributors remove, replace, or redact such information from datasets prior to upload. ● Notes: <ul style="list-style-type: none"> ○ Confidential or sensitive information refers to all identifiable information, such that re-identification of any subjects from the amalgamation of the information available from all of the materials is possible and can include: social security numbers; credit card numbers; medical record numbers; health plan numbers; other account numbers of individuals; or biometric identifiers (fingerprints, retina, voice print, DNA, etc.). ○ The repository defines a dataset as a collection of files, metadata, and ancillary content associated with the data. ○ Formats such as journal articles and conference papers resulting from research should be referred to the Institutional Repository. Users have the ability to link related content in different repositories through metadata fields. <p><i>Technical criteria:</i></p> <ul style="list-style-type: none"> ● UH IR: <ul style="list-style-type: none"> ☞ File uploads greater than 50 MB must be pre-approved by UH Libraries staff since large uploads require a mediated deposit process ☞ Projects (including those for individuals, classes, departments, and research centers) are subject to a 2 GB maximum limit ● UH Dataverse: <ul style="list-style-type: none"> ☞ File uploads can be up to 4GB per file.
<p>Audio/Video Repository</p>	<ul style="list-style-type: none"> ● Collections that meet the requirements for either UH Digital Collections OR the UH IR listed in this table, and that require any of the following features for improved user experience, expected use or reuse, or copyright status/contractual obligations: <ul style="list-style-type: none"> ☞ IIIF streaming media player, ability to embed media in other applications such as LibGuides, UH Libraries websites, or UH faculty research products ☞ Closed captions for video objects and associated transcripts for audio objects

	<ul style="list-style-type: none"> ∄ Download restrictions ∄ Ability to set different levels of access, including special access for certain users or user groups or IP addresses, either for the entire collection or for individual items within the collection
Third-party Repositories	<ul style="list-style-type: none"> ● Collections that are worthy of national recognition. They address critical elements including: <ul style="list-style-type: none"> ○ Aligning with popular research trends or issues ○ Aligning with one or more of the University of Houston’s strategic priority areas ○ Areas of local emphasis that would offer a unique contribution to the HathiTrust collection <p><i>Technical criteria:</i></p> <ul style="list-style-type: none"> ● Previously published in print ● Free of copyright restrictions ● Currently not available elsewhere on the web, including HathiTrust Digital Library, Internet Archive, and Google Books ● Fully described and marked up according to HathiTrust guidelines ● Complies with file format requirements of HathiTrust

Digital Collections Review Process

Focused on creating and implementing policies to increase national recognition around digital initiatives, the Digital Collections Management Committee (DCMC) has the explicit authority to:

- Approve or disapprove proposed projects using selection prioritization policies prior to projects entering any specific digital initiatives workflow
- Determine the repository for projects using digital deposit policies prior to projects entering any specific digital initiatives workflow

Digital collections are selected through the Call for Digital Projects (CFDP) process. During each cycle, DCMC will work closely with the Coordinator of Digital Projects, Coordinator of Metadata Services, Digital Operations Coordinator, and the Digitization Lab Manager to prepare the call and communicate selection decisions, including feedback when relevant. The committee will develop a mechanism for distributing the call and proposals will be submitted through an

online submission form (<https://uofh.sharepoint.com/sites/UHLibraries/dcmc>) found on the committee's intranet page. The number of projects selected during each cycle will depend on the extent of the collections and the resources available to digitize, describe, and make collections accessible. If a prior cycle's work has not been completed by the start of a new cycle, DCMC will work with the Coordinator of Digital Projects and the Digitization Services Coordinator to select the appropriate number of collections to include in the new cycle, including the possibility of skipping a call altogether.

In addition to selecting projects during the three cycles, DCMC will also evaluate time-sensitive proposals or those with external funding on a case-by-case basis and will make the appropriate changes to accommodate newly prioritized content. It may also solicit special CFDPs, that may fall outside of the regularly scheduled cycles, when needed.

The selection criteria for the CFDP includes:

- A. Collections that are worthy of national recognition. They address critical elements including:
 - a. Aligning with one or more of the University of Houston's strategic initiatives, including: Arts, Health, Energy, Athletics.
 - b. Making collections of unique or rare items of historical or cultural significance more accessible
 - c. Amplifying traditionally underrepresented voices in the existing collection
 - d. Complementing collections selected for commercial digital publication
 - e. Completing projects with defined external deadlines or with funding attached to the creation and dissemination of materials
- B. Collections that support the UH Division of Research's Research Thrusts
- C. Collections that support the research and student success objectives at the University of Houston. They address critical elements including:
 - a. Aligning selections with popular research trends or issues
 - b. Developing digital projects in conjunction with the Digital Research Commons and the Libraries' Sponsored Projects program
- D. Collections that fulfill important local needs. They address critical elements including:
 - a. Important local projects and university connections:
 - b. Complementing or completing existing digital collections (local or external)
 - c. Has potential for external funding

Reappraisal Guidelines

To better balance the research potential of collections with the necessary allocation of resources (space, staff, time, and conservation resources) for their care and preservation, DCMC should reappraise digital collections over time to determine whether some collections can no longer be supported. This should be an open, structured process conducted in consultation with content selectors and curators.

Policy Review

- Completed by DCMC: June 2, 2015
- Approved by Library Administration: September 23, 2015

Members of DCMC will review and update this document every two years.

Policy Revisions

- 2023-12-12: added language that refined the definition, purpose, and scope of Cougar ROAR and the A/V Repository.
- 2023-05-10: added storage limits for UH IR.
- 2021-11-03: added EDI-related language to selection criteria.
- 2020-08-19: added language that refined the definition, purpose, and scope of Cougar ROAR and the A/V Repository.
- 2018-01-26: added information and policies for UH Dataverse, HathiTrust, and the Call for Digital Projects process; revised policies for IR.
- 2015-07-30: revised commercial digital publishers section to state that DCMC reviews any proposal and gives recommendation to library administration. Also defined information that the Head of Special Collections should submit to DCMC with any proposals from a commercial digital publisher.
- 2015-07-28: migrated policy to UH Libraries Digital Initiatives Policies and Guidelines document. Eliminated “Related Policies and Guidelines” section because it is now part of the DIPG table of contents

References

- Amherst Library: <https://www.amherst.edu/library/services/digital/digitalcolldev>
- Carolina Digital Repository: <https://cdr.lib.unc.edu/scontent/aboutPages/policiesGuidelines.xml>
- Georgia Southern University: http://digitalcommons.georgiasouthern.edu/collection_development.html
- HathiTrust: http://www.hathitrust.org/take_down_policy
- Salt Lake Community College: <http://slccarchivesir.files.wordpress.com/2010/09/ir-collection-development-policy-013009.pdf>
- Society of American Archivists: <http://www2.archivists.org/sites/all/files/GuidelinesForReappraisalAndDeaccessioning-May2012.pdf>
- University of Vermont: <http://cdi.uvm.edu/about/colldev>