

DSPACE DIGITAL REPOSITORIES POLICIES

The purpose of UHM digital collections is to improve and stabilize access to digital material, including works of creative and scholarly output of value and potential interest to researchers. These guidelines are intended to apply to all digital repositories supported and maintained by the UHM Library, using the DSpace platform. Currently those repositories include ScholarSpace and eVols. In this document these digital repositories will be referred to as DDR = DSpace Digital Repositories.

CONTENT AND COLLECTION GUIDELINES

Characteristics of Content

1. Contributions to repositories should not be short-term; deposits are intended to be permanent contributions.
2. The work must be in digital form.
3. The work should be complete and ready for public dissemination before final submission.
4. DDRs can accept the following items:
 - a. Documents (e.g. articles, preprints, working papers, technical reports, conference papers)
 - b. Books
 - c. Theses and dissertations
 - d. Data sets
 - e. Computer programs
 - f. Visualizations, simulations, and other models
 - g. Multimedia publications
 - h. Learning objects
5. For DDRs, the UHM Library will accept all manner of digital formats, but will provide differing levels of support depending on the format type. See section “Access, Storage and Preservation” for additional information.

Additional Notes

1. Established Communities should strive for completeness in their collections. For example, if a work is part of a larger series or set of related works, these other works should also be contributed so that as full a set as possible is offered.
2. DDRs are not intended to be used for commercial purposes.
3. Digitization of the item must be complete. Portions of materials, such as a book page, will not be added to eVols, unless the item is significant in its present state to warrant such treatment.
4. Duplication within eVols or within ScholarSpace should be avoided. Use Item Mapping instead.

ESTABLISHING A COMMUNITY

Defining Communities and Collections

DSpace, the software underlying ScholarSpace and eVols, uses the following groupings to hierarchically organize content:

- **Communities** are top-level organizational units. In this document we also use the term Community to refer to the entity or person who is managing a Community.
- **Sub-Communities** are optional organizational units that can be established under Communities. Sub-Communities can themselves contain Sub-Communities.
- **Collections** are the organizational units where items/files (“Submissions”) are stored. Collections reside within Communities or Sub-Communities. Submissions must be placed within Collections, and cannot be placed directly into a Community or Sub-Community.

Community Start Up Procedures

1. Initial contact between the DDR staff and the prospective Community, and including the librarian liaison if applicable.
2. The DDR administrator demonstrates ScholarSpace and/or eVols to the prospective Community and provides information about the repository, policy documents, sample submission agreements, and workflow options.
3. Community chooses a coordinator to act as a liaison to DDR staff.
4. The community, in partnership with the DDR administrator, develops Community policy regarding submissions, access, content withdrawal, workflow, membership, and preservation strategies. A list of those members authorized to perform workflow steps should also be provided to the DDR administrator at this time.
5. DDR staff creates the Community, initial Collections and the corresponding interfaces in consultation with the Community coordinator. The DDR administrator then authorizes groups and individuals as outlined by the Community, and posts the final version of the Community policies, as well as written agreements between the Community and DDR on the Community webpage for future reference.

Community Responsibilities

The responsibilities taken on by a Community are to

1. Arrange for submission and description of content.
2. Decide policy regarding content to be submitted (within the appropriate DDR guidelines).
3. Make decisions about community and collection definitions and community membership.
4. Notify the DDR contact of organizational changes affecting submissions.
5. Understand and observe Institute policies relevant to the DDR, and educate community submitters regarding these policies.
6. Clear copyright for items submitted when copyright owner is other than author(s) or the UH.
7. Decide upon a submission workflow for each collection.

Community Rights

1. Decide who may submit content within the Community.
2. Limit access to content at the item level either to the UH only or to specific individuals or groups.
3. Receive a copy of submitted content upon request.
4. Remove items and collections (as outlined in "Withdrawal Policy").
5. Approve addition of or elimination of Sub-communities.
6. Customize interfaces to Community content.

DDR Responsibilities

At the UHM, the responsibilities taken on by DDR are to

1. retain and maintain submitted content.
2. distribute content according to Community decisions.
3. preserve content using accepted preservation techniques.
4. notify communities of significant changes to content, e.g. format migration.
5. if UH Manoa Libraries cease to support DDR, return collections to existing communities.

DDR Rights

DDR at the UHM Library retain the right to

1. redistribute, sell or amend metadata for items in DDRs.
2. de-accession items or Collections under certain circumstances - as outlined in "Withdrawal Policy."
3. refuse items or Collections not within the scope of the repository as defined by the Content Guidelines.
4. renegotiate terms of original agreement with Communities.
5. perform appraisal for long-term archiving when Communities cease to exist or within thirty years of the creation of a Collection.
6. move Collections to reflect current agreement between ScholarSpace or eVols and their Communities.
7. migrate items for presentation purposes or at the library's discretion.
8. set quotas (size of files, number of items) to determine what constitutes free service and after which point to charge a fee.
9. charge a fee for activities requiring extensive centralized support from the DDR (for example, for a large amount of de-accessioning)

SUBMISSION PROCESS

Each collection can have its own submission process. The community or collection liaison works with the ScholarSpace or eVols project to decide on the workflows they wish to use. DDR Communities will be asked to sign a Memorandum of Understanding with the UHM Library.

Authorization of Contributors

1. Communities determine contributors for each Collection

2. Authorization set out in Community policies

Submission Steps for Authorized Contributors

1. Register with ScholarSpace or eVols and edit his/her user profile
2. Post a submission to a specified Collection
3. Enter metadata for a submission, using standard forms with Community-specific default values
4. Include additional metadata with a submission
5. Bundle multiple files of various formats in a single submission
6. Allow the system to identify formats of submitted files, with user override
7. Grant a non-exclusive license to the UHM Library to distribute a submission in accordance with the specified distribution policy and translate it for the purposes of preservation (See note 1: Are the submitters granting permission to ScholarSpace or eVols, the UHM Library, or UHM?)
8. Receive updates if an item is rejected or at the end if the item gets final approval.
9. View past and pending submissions in a personalized area called My ScholarSpace or My eVols.

COPYRIGHT AND LICENSES – INTELLECTUAL PROPERTY RIGHTS (IPR)

Content Submitters Copyright

1. Although the author/copyright owner retains copyright for all works submitted, he/she should be willing and able to grant UHM Library the right to preserve and distribute the work via the digital repository.
2. Authors retain the copyright for all content posted in the repository. The author agreement specifies a nonexclusive right to use. This means the author is free to reuse the content elsewhere.

The UHM Library use the following license agreement in the DDR

Non-Exclusive Distribution License

In order for UHM Library to reproduce, translate and distribute your submission worldwide your agreement to the following terms is necessary. Please take a moment to read the terms of this license, fill in the information requested (and sign and submit this license to the DDR at _____.)

By signing and submitting this license, you (the author(s) or copyright owner) grants to the University of Hawaii at Manoa (UHM) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video.

You agree that UHM may, without changing the content, translate the submission to any medium or format for the purpose of preservation.

You also agree that UHM may keep more than one copy of this submission for purposes of security, back-up and preservation.

You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright.

If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant UHM the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission.

IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN UHM, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT.

UHM will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

The collection liaison is responsible for ensuring compliance with publisher copyright issues. However, the DDR project will provide information and resources to help with this assessment.

Orphan Works

Orphan works, described in the 2006 Copyright Office report on Orphan Works as "a term used to describe the situation where the owner of a copyrighted work cannot be identified and located by someone who wishes to make use of the work in a manner that requires permission of the copyright owner," may be added to ScholarSpace or eVols under the following conditions:

1. A good faith effort has been made to find and ask the copyright holder for permission to post [need to have some standardized definition of what a good faith effort is for UHM]
2. The orphan work will be removed from ScholarSpace or eVols if and when the copyright holder comes forward and objects to the work being online.

METADATA

Provenance Data

The DDR policy tracks the following provenance data:

1. Creation of item, Collection, or Community
2. Changes
 - a. Accessibility (see withdrawal policy)
 - b. Format
 - c. Organizational
 - i. Per item
 - ii. Per collection
 - iii. Per community
3. Metadata

4. Withdrawal of item, Collection, or Community

By “support” for a given metadata schema we mean that metadata can be entered into the DDR, stored in the database, indexed appropriately, and made searchable through the public User Interface. At the present time, this applies mainly to descriptive metadata, although as standards emerge it could also include technical, rights, preservation, structural and behavioral metadata.

Metadata Elements

The metadata elements in the table below are collected in the UH DDR for the following reasons:

1. to aid in the retrieval process
2. as a surrogate for the item (for instance, metadata harvesting by another system)
3. for use in later products (for instance, a bibliography in a particular discipline)

Metadata Element	Element Description	Policy
Contributor/Creator	An entity primarily responsible for or contributing to the making of the content of the resource	Required if available
Coverage	The extent or scope of the content of the resource	Not required
Date	A date associated with the lifecycle of the resource	Required
Description	An account of the content of the resource	Encouraged
Format	The physical or digital manifestation of the resource	System supplied
Identifier	An unambiguous reference to the resource within a given context	System supplied
Language	A language of the intellectual content of the resource	* Required (pull-down menu, including "non-text")

Publisher	An entity responsible for making the resource available	Not required
Relation	A reference to a related resource	Required if available
Rights	Information about rights held in and over the resource	Not required
Source	A reference to a resource from which the present resource is derived	Not required
Subject and Keywords	The topic of the content of the resource	Not required
Title	A name given to the resource	* Required
Type	The nature or genre of the content of the resource	Required

WORKFLOW

At the UHM Library the DDR staff recognizes that Communities have very different ideas for how material should be submitted to the repository, by whom, and with what restrictions. Who can deposit items? What type of items will they deposit? Who else needs to review, enhance, or approve the submission? To what collections can they deposit material? Who can see the items once deposited? All these issues are addressed by the Community representatives, working together with the Library's DDR user support staff, and are then modeled in a workflow for each collection to enforce their decisions. The system has the notion of "e-people" who have "roles" in the workflow of a particular Community in the context of a given collection. Individuals from the Community are registered with ScholarSpace or eVols, then assigned to appropriate roles. The following are typical Community representatives:

1. Submitter
2. Reviewer
3. Coordinator
4. Metadata Editor

These roles in the DDR support the workflow process: submitter, reviewer, coordinator, and metadata editor. An e-mail message is sent to each person at the appropriate step in the workflow, with authorizations set up in advance for each role.

Submitter permissions	<ul style="list-style-type: none"> - Can edit metadata for own submission - Can upload files for own submission - Cannot do anything once item is submitted
Reviewer permissions (optional)	<ul style="list-style-type: none"> - Can review content of all files submitted to collection - Can accept or reject all submissions to collection - Can send a message explaining decision - Rejection will stop submission - Acceptance will let submission go to next step - (Cannot edit metadata, or change files)
Coordinator permissions (optional)	<ul style="list-style-type: none"> - Can edit metadata of all submissions to collection - Can accept or reject all submissions to collection - Can send a message explaining rejection - Rejection will stop submission - Acceptance will move submission to next step
Metadata editor permissions (optional)	<ul style="list-style-type: none"> - Can edit metadata of all submissions to collection - Submission automatically becomes part of the DDR after this step - (Any approval would have happened before)

DDR staff, under the advice of the Cataloging Department, will work with Communities to identify unique metadata needs.

CONTRIBUTORS, USERS AND PRIVACY

Contributor Policies

1. Authorization to upload content into the repository, and responsibility for setting deposit policies will reside primarily within the registered Communities.
2. The Community retains the right to:
 - a. impose mandatory “embargo” periods that must expire before a digital object can be accessed by users.
 - b. limit access to content at the item level to the UHM Community.
 - c. remove items or Collections from any access under circumstances such as legal order, patenting guidelines, or publisher requirements.
3. Users will register with the ScholarSpace or eVols system and will be authenticated by the UHM system. Permissions will be assigned to users by each Collection Administrator.

4. Some items may require limitations to access. For example some theses be unavailable for a period of years. Other contributors may wish to limit access to their content.

Privacy Policies About Site Users

[need to determine whether there is such thing as an “DDR Privacy Policy” or if this is the policy. If so, should change DDR Privacy Policy to Privacy Policy for DDR.]

The DDR Privacy Policy at the UHM states that the University is committed to preserving privacy. The personal information the UHM receives through the DDR is used solely for purposes of the functioning of the system, and for the specific research purposes described below.

1. This system collects personal information from:
 - a. Users involved in the submission of DDR content and metadata
 - b. Users who subscribe to the DDR alerting service
2. Personal information collected by the DDR system will not be used for any commercial or philanthropic purpose not directly connected with or approved by the UHM.
3. We do not disclose information about your individual visits to our site, or personal information that you provide us, such as your name, address, email address, telephone number, etc. to any outside parties except when we believe, in good faith
 - a. that the law requires it, or
 - b. that disclosure is necessary to protect the rights and property of the DDR users.
4. Any DDR records used in a publicly accessible forum, such as demonstrations, presentations, or research papers, will be scrubbed of specific references to real people and personal information.

ACCESS, STORAGE AND PRESERVATION

ADA

Content should adhere to UHM ADA rules, and any applicable exemptions. [need to figure out who is responsible for compliance – the library or the contributor? And then add that responsibility to the rights/responsibilities sections, and possibly workflow section.]

Brief Overview of Formats

This next two sections, “Preservation Formats” and “Levels of Preservation,” discuss management of different file types, in technical detail. Briefly and in plain language:

With regard to file types, key points can be summarized as follows:

1. All file types are accepted and will be maintained through bit-level preservation.
2. However, not all file types are displayable within our digital repositories, and not all file types can be supported, i.e. made usable in the future.

See “Preservation Formats” and “Levels of Preservation” for details.

Preservation Formats

Digital preservation is an evolving practice and standards are not yet in place. However, we commit to tracking the development of standards and issuing advisories to Communities who are submitting to the repository. The UHM Library implementation of the DDR systems attempts to support as many file formats as possible. This system identifies two levels of digital preservation: bit preservation, and functional preservation.

1. **Bit preservation** ensures that a file remains exactly the same over time - not a single bit is changed - while the physical media evolve around it.
2. **Functional preservation** goes further: the file does change over time so that the material continues to be immediately usable in the same way it was originally while the digital formats (and the physical media) evolve over time. Some file formats can be functionally preserved using straightforward format migration (e.g. TIFF images or XML documents). Other formats are proprietary, or for other reasons are much harder to preserve functionally.

Levels of Preservation

At UHM Library, for the time being, we acknowledge the fact that the formats in which people create their content are not something we can predict or control. People use the tools that are best for their purposes, and we will get whatever formats those tools produce. Because of this we've defined three levels of preservation for a given format: supported, known, or unsupported.

1. **Supported:** We fully support the format and preserve it using either format migration or emulation techniques.
2. **Known:** We can recognize the format, but cannot guarantee full support.
3. **Unsupported:** We cannot recognize a format; these will be listed as "application/octet-stream", aka Unknown.

When a file is uploaded to ScholarSpace or eVols, we assign it one of those three categories. For all three levels we will do bit-level preservation so that "digital archaeologists" of the future will have the raw material to work with if the material proves to be worth that effort.

Put simply, UHM Library' policy for file formats is:

1. Everything put into ScholarSpace or eVols will be retrievable.
2. We will recognize as many files' formats as possible.
3. We will support as many known file formats as possible.

By "support", we mean "make usable in the future, using whatever combination of techniques (such as migration, emulation, etc.) is appropriate given the context of need." We can't predict which services will be necessary down the road, so we'll continually monitor formats and techniques to ensure we can accommodate needs as they arise.

In the meantime, we can choose to "support" a format if we can gather enough documentation to capture how the format works. In particular, we collect file specifications, descriptions, and code samples, and make those available in the Format Reference Collection as listed in the table below. Unfortunately, this means that proprietary formats for which these materials are not publicly available cannot be supported in DDR. We will still preserve these files, and in cases

where those formats are native to tools supported by UHM 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."

Format Reference Collection for UHM installation of DDR

In the table below, MIME type is the Multipurpose Internet Mail Extensions (MIME) type identifier. Description is what most people use as the name for the format. Extensions are typical file name extensions (the part after the dot, e.g. the extension for "index.html" is "html"). These are not case-sensitive in the DDR, so either "sample.XML" or "sample.xml" will be recognized as XML. Level is the DDR's support level for each format:

1. **Supported:** We fully support the format and preserve it using either format migration or emulation techniques.
2. **Known:** We can recognize the format, but cannot guarantee full support.
3. **Unsupported:** We cannot recognize a format; these will be listed as "application/octet-stream", aka Unknown.

MIME type	Description	Extensions	Level
Application/marc	MARC	marc, mrc	supported
Application/mathematica	Mathematica	ma	known
Application/msword	Microsoft Word	doc	known
Application/octet-stream	Unknown	(anything not listed)	unsupported
Application/pdf	Adobe PDF	pdf	supported
Application/postscript	Postscript	ps, eps, ai	supported
Application/sgml	SGML	sgm, sgml	known
Application/vnd.ms-excel	Microsoft Excel	xls	known
Application/vnd.ms-powerpoint	Microsoft Powerpoint	ppt	known

Application/vnd.ms-project	Microsoft Project	mpp, mpx, mpd	known
Application/vnd.visio	Microsoft Visio	vsd	known
Application/wordperfect5.1	WordPerfect	wpd	known
Application/x-dvi	TeXdvi	dvi	known
Application/x-filemaker	FMP3	fm	known
Application/x-latex	LateX	latex	known
Application/x-photoshop	Photoshop	psd, pdd	known
Application/x-tex	TeX	tex	known
audio/x-aiff	AIFF	aiff, aif, aifc	supported
audio/basic	audio/basic	au, snd	known
audio/x-mpeg	MPEG Audio	mpa, abs, mpeg	known
audio/x-pn-realaudio	RealAudio	ra, ram	known
audio/x-wav	WAV	wav	known
image/gif	GIF	gif	supported
image/jpeg	JPEG	jpeg, jpg	supported
image/png	PNG	png	supported
image/tiff	TIFF	tiff, tif	supported
image/x-ms-bmp	BMP	bmp	known
image/x-photo-cd	Photo CD	pcd	known
text/html	HTML	html, htm	supported
text/plain	Text	txt	supported
text/richtext	Rich Text Format	rtf	supported

text/xml	XML	xml	supported
video/mpeg	MPEG	mpeg, mpg, mpe	known
video/quicktime	Video Quicktime	mov, qt	known

Persistent Identifiers

The CNRI Handle System is used to assign persistent identifiers to all material posted to the repository. These identifiers are resolvable in perpetuity, and will remain valid even if content is migrated to a new system. This allows documents in the DDR to be properly and effectively cited in other research.

Withdrawal of Items

1. UHM Library foresees times when it may be necessary to remove items from the repository. It has been decided that under some circumstances items will be removed from view, but to avoid loss of the historical record, all such transactions will be traced in the form of a note in the field of the Dublin Core record. The content of the note should be one of the following:
 - a. " removed from view at request of the author"
 - b. " removed from view at UHM's discretion"
 - c. " removed from view at UHM Library's discretion"
 - d. " removed from view by legal order"
2. If the intent is to remove all records and links to an item from the collection, it can be deleted upon request.
3. Since any DDR item that has existed at some time may have been cited, we will always supply a "tombstone" stating that the item was withdrawn.

Free and Fee Access

1. Initially DDR will offer Core Services at no charge to Community members and consumers of the content. Two distinct, but interconnected service elements are
 - a. Interactive Services. These offer a fully functional system that allows DDR Community members and consumers of the content to accomplish all tasks necessary to submit and access items in the repository as applicable.
 - b. Operational Services. These host and preserve faculty materials, establish and deliver ongoing support for the DDR Communities, respond to customer inquiries, and supply system monitoring, back up, and recovery.
2. It is conceivable that the DDR Communities or individual submitters may put extraordinary demands on the service such as sizeable storage requirements or assistance with specialized metadata creation. In the future, the UHM Library may offer Premium Services to ensure that DDR offer a full set of resources to meet contributors' and content users' needs and to manage the impact of these exceptional resource requirements on Library staff and DDR resources. UHM Library reserves the right to introduce Premium

Services fees as needed to aid in cost recovery. Possible fee-based Premium Services include the following:

- a. E-Conversion Services: Creation of digital content from non-digital materials and custom, on-demand transformation of materials from one format to another.
- b. Metadata Services: Needs assessments, feasibility studies, advice on appropriate taxonomies, metadata crosswalks, metadata creation and support services, etc.
- c. Custom Repository Services: Expansion of standard ScholarSpace or eVols storage allocations to meet Community or individual's requirements that exceed normal limits
- d. User Reporting Services: Research alert services, targeted notification services, hot topic citations, and custom reporting services

Backup and Recovery

1. The system may be unavailable on occasion for update purposes. However, the DDR system administrators will perform maintenance at times least likely to cause a disruption of services.
2. The DDR administrative team will provide regularly-scheduled backup with off-site storage to guarantee backup and recovery.