

Welcome To The Connecticut Digital Archive

*Long Term Preservation Services for
Nonprofit Connecticut Based Institutions*

Topics

- What are the technologies of the CTDA?
- Namespaces
- Organizing Your Content
- The Sandbox
- Your Profile
- Top Level Navigation
- Managing content
- Users
- Resources

What is the CTDA?

- The CTDA is a digital repository.
- The CTDA infrastructure is built on the Open Archival Information System (OAIS) developed by NASA and now an ISO standard (14721:2003)
- OAIS is a conceptual model of an information object that is self-contained and self-describing. An information object is a set of data elements combined into an internally coherent package that is understood in the electronic and web environment.

Primary Technologies of the CTDA

- FedoraCommons (<http://www.fedora-commons.org/about>)
 - Not a database but a conceptual framework
 - Not a software application but a generic repository foundation upon which many kinds of applications can be created
 - Not an integrated solution or information silo but a basis for software systems to manage digital information
- Islandora (<http://islandora.ca/about>)
 - Administrative, Discovery and Presentation layer build on Drupal
- Handle System (<http://www.handle.net/>)
 - Persistent and unique identifier generation system and lives in the MODS metadata



The CTDA Stack



1. Create, Manage, Share
Information

2. Manage and Preserve
Repository and Archive Application, Storage
Preservation

3. Store
Storage Devices

Talks to

Handle System®

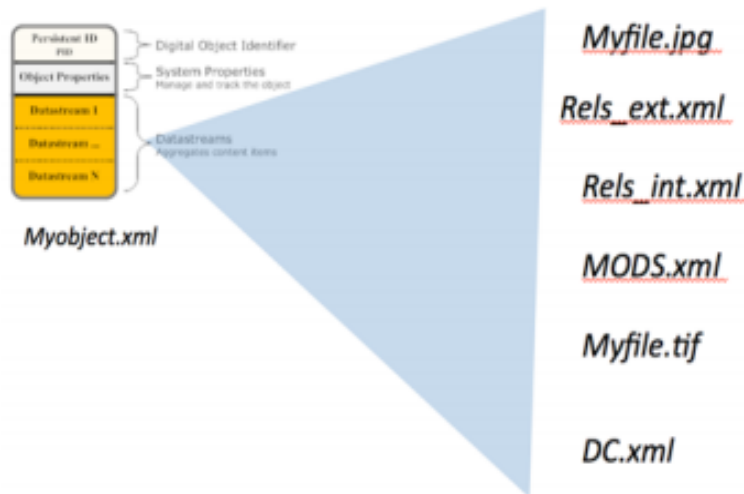
*System that automatically creates persistent
identifiers for citation purposes*

A Brief Introduction to FEDORA

Fedora Digital Objects: Information Packets

Packages and Files

ALL information about an object, including the primary digital content is contained in a set of files called **datastreams** that are organized into an “information package” called a **repository object**



OAIS information package

Fedora Objects Requirements

- All Fedora Objects must have:
 - A local and internal persistent identifier called the PID. This is not the identifier used for citation purposes which is the handle.
 - Datastream(s) for Object Properties that tells the system where the Fedora Object belongs in relation to other Fedora Objects in the repository.
 - Dublin Core datastream
- The CTDA uses two types of Fedora Objects: Container (or collection) and Content Fedora Digital Objects.

Container or Collection Objects

- The container object is the simplest type of Fedora object in that there is just metadata about that collection in various data streams, an internal identifier (PID), and the objects “coordinates”, the rels-ext.
- How and why do you describe a collection? It’s not necessary. However, you might want to designate the name of a collection and provide a description of what’s in the collection for instance.
- Container Objects may contain other containers in an infinite regression.
- Each object has data that positions it in relation to other objects or provides analogously the object’s coordinates in the repository, the rels-ext.

Groton Public Library – 180002:GPL001

ID	LABEL	TYPE	MIME TYPE	SIZE
RELS-EXT	Fedora Object to Object Relationship Metadata.	Inline XML	application/rdf+xml	992 B
COLLECTION_POLICY	Collection policy	Inline XML	text/xml	816 B
MODS	MODS Record	Managed	text/xml	923 B
DC	DC Record	Inline XML	text/xml	536 B
TN	TN	Managed	image/jpeg	17.6 Ki
POLICY	XACML Policy Stream	Inline XML	text/xml	4.06 Ki

Content Object

- The content object has more data streams in that it has all that the container object does in addition to the digital asset (archival master) and its derivatives one metadata record describes one primary content data stream that many have a number of derivatives.
- This content object is independently discoverable, but the individual parts are not.
- All individual parts are independently addressable.

Branford House, Plant's Residence, Eastern P			
ID	LABEL	TYPE	MIME TYPE
RELS-EXT	Fedora Object to Object Relationship Metadata.	Inline XML	application/rdf+xml
MODS	ck141A.xml	Managed	text/xml
DC	DC Record	Inline XML	text/xml
OBJ	ck141A.tif	Managed	image/tiff
TECHMD	TECHMD	Managed	text/xml
TN	Thumbnail	Managed	image/jpeg
JPG	Medium sized JPEG	Managed	image/jpeg
JP2	JPEG 2000	Managed	image/jp2



A Brief Introduction to Islandora

Islandora the framework...

- Allows you to manage and display content.
- Is a set of Drupal modules that facilitate interaction with Fedora through pages and forms.
- Basis for Management and Presentation interfaces to the Fedora repository
- Can be used for multiple sites. The CTDA has a multisite implementation. This means that there is ONE Fedora repository but multiple Islandora/Drupal sites that interact with that one repository.

Manage CTDA Collections

Connecticut State Library

UNIVERSITY LIBRARIES

Archives & Special Collections at the Thomas J. Dodd Research Center

Trinity College Library

The Connecticut Digital Archive Collections

Islandora: Architecture Basics

Content Models

- ***Content Models*** or ***Islandora Content Model Digital Objects*** determine how repository Objects are created and managed in the repository and displayed to the user.
- They are **for the most part** associated with specific file formats because these Content Models tell the system if and which...
 - Viewer to call
 - Derivatives to create
 - Derivatives that are available for download
 - Technical metadata is extracted
 - Object properties are created
 - Specific relationships to other Fedora objects to be created
 - Specific access policies
 - *And so much more...*

Content Models & Formats

- **Collection:** Used for container objects (no format as no content is involved)
- **Large Image:** tiff, jp2 (Viewer has zoom functionality)
- **Basic Image:** jpg, png, bmp, gif (Simple viewer with no zoom functionality)
- **PDF:** PDF files (PDF.js viewer replicates PDF viewer for user in display)
- **Audio:** wav, mp3
- **Video:** ogg, mp4, mov, qt, m4v, avi
- **Book and page:** A specialized set of content models for ordered sets of digital files. The book is a container. Pages need to be tiff/tif or jp2 format. There is also a book batch import that can take a PDF file.

Content Models & Formats

- **Newspaper, issue and page:** A specialized set of content models for serial publications or regularly recurring ordered sets of digital files. The newspaper and issue are containers. Pages need to be tiff/tif, jp2, or jpg format. There is also a batch import that can take a PDF.
- **Compound Object:** A specialized set of content models for sets of digital files that combine other Objects in an explicit relationship (i.e. front and back of a postcard). All child objects have to be ingested but can be “hidden”.
- **Citation:** A specialized content model that will record a bibliographic citation according to a specific citation style guide such as APA or Chicago Manual of Style. Batch import can be used for PubMed articles. This is particularly useful for faculty citations to materials that cannot be added to a publicly available repository.

Content Models & Formats

- **Document:** Used for common documents (odt, sxw, rtf, doc, wpd, txt, html, ods, sxc, xls, csv, tsv, odp, sxi, ppt, pdf, docx, xlsx, pptx). Batch import will work for all formats.
- **Thesis:** Used for theses and dissertations. This includes the ability to embargo objects or datastreams.
- **WebARChive:** Used for the web archive format (warc), a format revision of the Internet Archive's ARC file format (warc). Upon ingest can also include a screen shot of the web site (as png or jpg) and accompanying txt file.
- **Entity/Person:** Used to add an object about a person such as a profile of a faculty member. This doesn't involve content and is a specialized container object – specialized in terms of how it is displayed and the metadata used to describe the person (EAC-CPF or MADS).
- **Entity/Place:** Used to add an object about a place such as a profile of a city. This doesn't involve content and is a specialized container object – specialized in terms of how it is displayed and the metadata used to describe the person (EAC-CPF or MADS).

Content Models & Formats

- **Entity/Organization:** Used to add an object about an organization such as a profile of a vessel. This doesn't involve content and is a specialized container object – specialized in terms of how it is displayed and the metadata used to describe the person (EAC-CPF or MADS).
- **Entity/Event:** Used to add an object about an such as a profile of a hurricane. This doesn't involve content and is a specialized container object – specialized in terms of how it is displayed and the metadata used to describe the person (EAC-CPF or MADS).
- **Finding Aid:** Used to add archival finding aids (EAD xml files)
- **Binary Object:** A specialized content model for alternative formats not listed above

A Brief Word on Organizing Your Content In The Repository

Organizing Your Content

- Learning how to add and manage content using Islandora is like riding a bike. Once you learn it, you have it.
- Organizing your content is different and requires strategies
 - Content can be organized differently.
 - Collections can be nested or not
 - Digital objects can be “shared”
 - Digital objects can be “migrated”
 - There’s not one way to add content.
 - Different content models can be applied to content.
 - You don’t have to apply the same strategy to everything.
 - You can change your mind.

Information Organization

- How you organize your content depends on:
 - Your users
 - How you would like your users to discover your digital assets
 - How you think users will discover and access your digital assets
- There is not one way to organize all of your content. You can use multiple methods. You can even change your mind.
- Before you add and manage content, it is worth developing an initial method of how you'll organize content and even how you will describe it.

Managing Content From Different Institutions in one Digital Repository

How do we ensure content is
associated with the correct
institution?

Namespaces

The CTDA is made up of several partners who all have unique content but share the same digital repository. The CTDA uses **NAMESPACES** or a series of digits to identify content from an institution and sometimes a particular set of content within that institution. One institution can have 1+ namespaces.

For a list of our most current namespaces, see <http://ctdigitalarchive.org/about/technology/namespaces/>.

Why Are Namespaces Important?

- Distinguish your content from other institutions or other sets of content within your institution
- Help you manage your content more appropriately (i.e. an institution can have a range of namespaces that represent different sets of data such as research data, electronic regulations, etc.)
- Add a new container/collection object
 - You **MUST** create a PID for all container objects. PIDs are your namespace:alphanumericString with NO SPACES or special characters. However, you can use underscores and hyphens.
 - Remember to “Uncheck” “Inherit Collection Policy” when creating a container/collection object

[Home](#)
[Browse repository](#)

Collection PID

Unique PID for this collection. Leave blank to use the default.
PIDs take the general form of **namespace:collection** (e.g., islandora:pamphlets)

☐ Inherit collection policy?

COLLECTION POLICY

Children namespace

Namespace for objects with selected content models.

Allowable content models

- ☐ islandora:pageCModel ~ Islandora Page Content Model
- ☐ islandora:sp_large_image_cmodel ~ Islandora Large Image Content Model
- ☐ islandora:bookCModel ~ Islandora Internet Archive Book Content Model
- ☐ islandora:sp_basic_image ~ Islandora Basic Image Content Model
- ☐ islandora:sp-audioCModel ~ Islandora Audio Content Model
- ☐ islandora:collectionCModel ~ Islandora Collection Content Model
- ☐ islandora:sp_videoCModel ~ Islandora Video Content Model
- ☐ islandora:compoundCModel ~ Islandora Compound Object Content Model
- ☐ islandora:binaryObjectCModel ~ Islandora Binary Object Content Model
- ☐ islandora:sp_pdf ~ Islandora PDF Content Model
- ☐ islandora:newspaperPageCModel ~ Islandora Newspaper Page Content Model
- ☐ islandora:newspaperIssueCModel ~ Islandora Newspaper Issue Content Model
- ☐ islandora:newspaperCModel ~ Islandora Newspaper Content Model

Content models describe the behaviours of objects with which they are associated.

Inherit XACML policy from

Previous Next

CTDA Sandbox

<http://ctda-stage.lib.uconn.edu/>

Your Profile

- You can access your profile in one of 2 ways.
 - After you log in
 - By clicking on your username in the upper right corner
- To make changes to your profile, click “Edit”
 - You can change your username and password
 - You can also upload a profile picture
- Remember to save all changes
 - Click on “Save” at the bottom
- If you forget your password, just email ctda@uconn.edu



jeustis

View

Edit

History

Member for
8 months 4 days

Username *

jeustis

Spaces are allowed; punctuation is not allowed

Current password

Enter your current password to change the E-n

E-mail address *

jennifer.eustis@lib.uconn.edu

A valid e-mail address. All e-mails from the sy:

Password

Pas:

Confirm password

To change the current user password, enter the

Top Level Navigation

- Getting to Collections
 - Click on “CTDA Stage”
 - Or search for something
- If you click on CTDA Stage, to see all the test collections, click on “Test Collections”
- You can use the “next”, “last”, etc. on the right hand side to navigate all the pages of Container Objects (for the most part these will be container Objects and not Content Objects)
- To see what’s “in” one of these Container Objects, just click on any one.

CTDA Stage

Manage & Browse Digital Collections

This site provides CTDA participants with an opportunity to ex environment. Participants are encouraged to create collections Since this is a staging and testing site, content from the site w prepare and test collections and objects prior to ingest into a p

Browse the Test Collections

1 2 3 next > last »



Create A New Collection (Container Object)

- No matter where you are in the system, it is the same process to create a Container Object. How does CTDA know where to create your Container Object? **You have to go to that location and then click manage.**
- Create a collection
 - Click Manage
 - Click “Add an object to this Collection”
 - Select “Islandora Collection Content Model”
 - Click Next
 - Fill out the Collection PID
 - **UNCHECK INHERIT COLLECTION POLICY – SEE**

Home > Digital Collections

My First Collection

View

Manage

[+ Add an object to this Collection](#)

Total members: 35

Select a Content Model to Ingest

- Islandora Collection Content Model
- Islandora Collection Content Model
- Islandora Large Image Content Model

Container Object Continued

- The Collection PID **MUST** have your institution's namespace colon and then any alphanumeric string
 - PIDs cannot contain any spaces or unusual characters such as slashes, question marks or brackets
 - PIDs can contain underscores and hyphens
- **UNCHECK** "Inherit collection policy"
- Leave the default for Inherit XACML policy from.

Collection PID

Unique PID for this collection. Leave blank to use the default.
PIDs take the general form of **namespace:collection** (e.g., islandora:pamphlets)

☐ Inherit collection policy?

COLLECTION POLICY

Children namespace

Namespace for objects with selected content models.

Allowable content models

- ☐ islandora:pageCModel ~ Islandora Page Content Model
- ☐ islandora:sp_large_image_cmodel ~ Islandora Large Image Content Model
- ☐ islandora:bookCModel ~ Islandora Internet Archive Book Content Model
- ☐ islandora:sp_basic_image ~ Islandora Basic Image Content Model
- ☐ islandora:sp-audioCModel ~ Islandora Audio Content Model
- ☐ islandora:collectionCModel ~ Islandora Collection Content Model
- ☐ islandora:sp_videoCModel ~ Islandora Video Content Model
- ☐ islandora:compoundCModel ~ Islandora Compound Object Content Model
- ☐ islandora:binaryObjectCModel ~ Islandora Binary Object Content Model
- ☐ islandora:sp_pdf ~ Islandora PDF Content Model
- ☐ islandora:newspaperPageCModel ~ Islandora Newspaper Page Content Model
- ☐ islandora:newspaperIssueCModel ~ Islandora Newspaper Issue Content Model
- ☐ islandora:newspaperCModel ~ Islandora Newspaper Content Model

Content models describe the behaviours of objects with which they are associated.

Inherit XACML policy from

None ▼

Previous

Next

Container Object Continued

- Select the content models that are appropriate for this collection
 - You want to create a collection to only ingest tiff images. Select Islandora Large Image Content Model
 - You want to create a collection to only ingest a book. Select Islandora Page Content Model and Islandora Internet Archive Book Content Model
 - You want to create a collection where you can ingest different formats. Select all the appropriate content models.
- Click “Next”
- Fill out the metadata form
- Click “Ingest” at the bottom of the form

the name of the organization

ous

Ingest

Allowable content models

- ☐ islandora:pageCModel ~ Islandora Page Content Model
 - ☐ islandora:sp_large_image_cmodel ~ Islandora Large Image Content Model
 - ☐ islandora:bookCModel ~ Islandora Internet Archive Book Content Model
 - ☐ islandora:sp_basic_image ~ Islandora Basic Image Content Model
 - ☐ islandora:collectionCModel ~ Islandora Collection Content Model
 - ☐ islandora:compoundCModel ~ Islandora Compound Object Content Model
 - ☐ islandora:sp_pdf ~ Islandora PDF Content Model
 - ☐ islandora:sp-audioCModel ~ Islandora Audio Content Model
 - ☐ islandora:newspaperCModel ~ Islandora Newspaper Content Model
 - ☐ islandora:newspaperPageCModel ~ Islandora Newspaper Page Content Model
 - ☐ islandora:sp_web_archive ~ Islandora Web ARChive Content Model
 - ☐ islandora:sp_videoCModel ~ Islandora Video Content Model
 - ☐ islandora:newspaperIssueCModel ~ Islandora Newspaper Issue Content Model
 - ☐ islandora:OACCModel ~ Islandora Annotation CModel
 - ☐ islandora:personCModel ~ Islandora Person Content Model
 - ☐ islandora:organizationCModel ~ Islandora Organization Content Model
 - ☐ islandora:entityCModel ~ Islandora Entity Content Model
 - ☐ islandora:eventCModel ~ Islandora Event Content Model
 - ☐ islandora:placeCModel ~ Islandora Place Content Model
 - ☐ islandora:binaryObjectCModel ~ Islandora Binary Object Content Model
- Content models describe the behaviours of objects with which they are associated.

Previous

Next

Add 1 Digital Asset (Your Content)

- Two options to add content: one at a time or by batch import.
Remember that your collection has to be set up with the appropriate content models. If you want to add a PDF and you didn't select that content model then you have to add that Content Model in Manage, then Collections and Collection Policies.
- Adding one digital asset at a time (Single upload)
 - Click "Manage"
 - Click "Add an object to this Collection"
 - If you set your Container Object for several Content models, you will need to select a Content Model to Ingest (i.e. you need to tell the system the format of your content, large image, basic image, PDF, audio, video). Once selected click "Next"
 - You might have several custom Metadata forms. If so, then select a Metadata entry form (default Islandora form, CHI, or CTDA)
 - Fill out the form about your content and click "Next"
 - Choose your file
 - You can click "Upload" and then "Ingest" or just "Ingest". There is no difference.

Batch Import Your Content

- Adding 1+ digital assets with or without associated metadata (Batch import)
 - All metadata files must be in XML and follow the CTDA MODS Implementation Guidelines (<http://ctdigitalarchive.org/resources-for-participants/>)
 - All XML files **MUST** be valid and well-formed
 - For content and their associated metadata files, the content file and MODS xml file **MUST** have the ***exact*** file name (i.e. myTiffimage01.tiff, myTiffimage.xml)
 - All files **MUST** be in a zipped folder.
- Book batch: This works to ingest 1 book and its pages. There's the MODS.xml file and each folder has the page number and in that folder is the page.

5	7/20/2015 5:52 PM	File folder
6	7/20/2015 5:52 PM	File folder
7	7/20/2015 5:52 PM	File folder
8	7/20/2015 5:52 PM	File folder
9	7/20/2015 5:52 PM	File folder
10	7/20/2015 5:52 PM	File folder
11	7/20/2015 5:52 PM	File folder
12	7/20/2015 5:52 PM	File folder
13	7/20/2015 5:52 PM	File folder
14	7/20/2015 5:52 PM	File folder
15	7/20/2015 5:52 PM	File folder
16	7/20/2015 5:52 PM	File folder
17	7/20/2015 5:52 PM	File folder
18	7/20/2015 5:52 PM	File folder
19	7/20/2015 5:52 PM	File folder
20	7/20/2015 5:52 PM	File folder
21	7/20/2015 5:52 PM	File folder
22	7/20/2015 5:52 PM	File folder
23	7/20/2015 5:52 PM	File folder
24	7/20/2015 5:52 PM	File folder
MODS.xml	7/20/2015 5:34 PM	XML File

Adding A Book

- A Book is essentially a container that has pages and is done in 2 steps unless you do book batch
 - Click Manage and then “Add an object to this collection”
 - Select “Islandora Internet Archive Book Content Model” and click Next
 - Describe the book using the form and click Next
 - You can add the PDF of the book pages. If you do this, you can select the Page Image Settings, Resolution and Page OCR/Text Settings. Choose and upload your file. Select your settings. Click “Ingest”
 - If you don’t have a PDF, then just select the language and click “Ingest”
- Book batch: This works to ingest 1 book and its pages. There’s the MODS.xml file and each folder has the page number and in that folder is the page.

PDF File

No file chosen

You can optionally upload a PDF file, from which the pages of the document will be generated.

PAGE IMAGE SETTINGS

Image Format

- ☐ 8-bit grayscale
- ☐ 12-bit RGB
- ☐ 24-bit RGB
- ☐ 48-bit RGB
- ☒ 32-bit CMYK
- ☐ 64-bit CMYK

You can choose the format for the page images extracted from the uploaded PDF. We recommend 32-bit CMYK for most cases.

Resolution (DPI)

- ☐ 72
- ☐ 96
- ☐ 150
- ☒ 300
- ☐ 600

The higher the resolution selected, the more legible the text will be, but be aware it takes longer to process higher resolution images, and they will consume more space.

PAGE OCR / TEXT SETTINGS

Extract Text

- ☐ Do not extract text.
- ☐ Perform OCR on the PDF.
- ☒ Extract text from the PDF.

Extracted text is used to aid in the discovery of this object when searching. If the uploaded PDF contains text and not just images of text, we recommend you try to extract

Language

English ▼

Please select the language the page is written in.

	FILE NAME	FILE TYPE
6	7/20/2015 5:52 PM	File folder
7	7/20/2015 5:52 PM	File folder
8	7/20/2015 5:52 PM	File folder
9	7/20/2015 5:52 PM	File folder
10	7/20/2015 5:52 PM	File folder
11	7/20/2015 5:52 PM	File folder
12	7/20/2015 5:52 PM	File folder
13	7/20/2015 5:52 PM	File folder
14	7/20/2015 5:52 PM	File folder
15	7/20/2015 5:52 PM	File folder
16	7/20/2015 5:52 PM	File folder
17	7/20/2015 5:52 PM	File folder
18	7/20/2015 5:52 PM	File folder
19	7/20/2015 5:52 PM	File folder
20	7/20/2015 5:52 PM	File folder
21	7/20/2015 5:52 PM	File folder
22	7/20/2015 5:52 PM	File folder
23	7/20/2015 5:52 PM	File folder
24	7/20/2015 5:52 PM	File folder
MODS.xml	7/20/2015 5:34 PM	XML File

Adding A Newspaper

- Newspaper is a 3 step process. There is the newspaper, the issue and the pages of the issue.
 - Click Manage and “Add an object to this collection” and click Next
 - Select “Islandora Newspaper Content Model”
 - Describe the newspaper in the form and click “Ingest”. This will create a “container” for the issues of this newspaper
 - In this newspaper, click Manage and “Add Issue”. Describe the issue and click Next.
 - You can add the PDF of the issue pages. If you do this, you can select the Page Image Settings, Resolution and Page OCR/Text Settings. Choose and upload your file. Select your settings. Click “Ingest”
 - If you don’t have a PDF, then just select the language and click “Ingest”
- Batch import pages of an issue: If you didn’t upload a PDF and have pages you want to batch import, go to the issue and click Manage. Go to the tab called “Issue” which opens a new window. Click “Add zipped pages”. Select your language and zip file to upload. Click upload and then “Add files”

PDF File

No file chosen

You can optionally upload a PDF file, from which the pages of the document will be generated.

PAGE IMAGE SETTINGS

Image Format

- ☐ 8-bit grayscale
- ☐ 12-bit RGB
- ☐ 24-bit RGB
- ☐ 48-bit RGB
- ☒ 32-bit CMYK
- ☐ 64-bit CMYK

You can choose the format for the page images extracted from the uploaded PDF. We recommend 32-bit CMYK for most cases.

Resolution (DPI)

- ☐ 72
- ☐ 96
- ☐ 150
- ☒ 300
- ☐ 600

The higher the resolution selected, the more legible the text will be, but be aware it takes longer to process higher resolution images, and they will consume more space.

PAGE OCR / TEXT SETTINGS

Extract Text

- ☐ Do not extract text.
- ☐ Perform OCR on the PDF.
- ☒ Extract text from the PDF.

Extracted text is used to aid in the discovery of this object when searching. If the uploaded PDF contains text and not just images of text, we recommend you try to extract

Language

Please select the language the page is written in.

[+ Add Page](#) [+ Add Zipped Pages](#)

Create PDF

Perform OCR

Create Images

Reorder Pages

Delete Pages

You must have both **ImageMagick** and **GhostScript** installed. This will also update the PDF datastreams for each Page

DPI – Dots Per Inch

Set the DPI for the generated PDF. The higher the resolution, the more legible the text will be, but be aware it takes longer to process higher resolution images, and they will consume more space.

Adding A Compound Object

- Compound objects are parent/child relationships
- All of the child digital assets have to be uploaded before creating the compound object. You can put the child objects in a collection of their own if you want
- Note the PIDs of the digital assets that are to be children – you’ll need to enter these PIDs later
- Create the parent.
 - Click Manage and “Add an object to this collection”
 - Select “Islandora Compound Object Content Model” and click Next
 - Describe the “parent” in the form and click “Ingest”
 - Click Manage again and click on the tab “Compound”
 - In the new window, enter the 1st PID of the child digital asset and click “Save”. Repeat for all children.

ADD CHILD OBJECTS

Add child objects as part of this compound object

Child Object Pid/Label

ASSOCIATE WITH PARENT OBJECT

Add this object to a parent object

Parent Object Pid/Label

Save

Description

Adding A Web Archive File

- The Islandora Web Archive Content accepts warc file formats, a screen shot of that website and an associated text file.
- The text file needs to have the extension .txt.
- The Screen shot can be either an image that is the format jpg/jpeg or png.
- The warc format specifies a method for combining multiple digital resources into an aggregate archival file together with related information. It is a revision of the Internet Archives' ARC file format. More information can be found at:
<http://www.digitalpreservation.gov/formats/fdd/fdd000236.shtml>
- If you have a warc file
 - Click Manage and “Add an object to this collection”
 - Select “Islandora WebARChive Content Model” and click Next
 - Describe the file and click “Next”
 - Select your warc file and click “Upload”
 - If you want to add a screen shot, click the box where you can select your image and upload. Same thing for your text file.

Warc *

No file chosen

Select warc to upload.
Files must be less than 128 MB.
Allowed file types: warc.

☐ Upload a screenshot?

☐ Upload a **Text file**

Previous

Ingest

Adding A Document

- The Islandora Document Content Model is used for most documents such as word, text files, or pdf.
- Click Manage and “Add an object to this collection”
- Select Islandora Document Content Model and Next
- Describe the document and click “Next”
- Choose your
 - If you have a PDF, odt, sxw, doc, docx, wpd, txt, html, ods, sxc, xls, csv, tsv, odp, sxi, ppt, pptx, xlsx file type
 - Click Manage and “Add an object to this collection”
 - Select “Islandora Document Content Model” and click Next
 - Describe the file and click “Next”
 - Select your file and click “Upload”

Document File *

Choose File

No file chosen

Upload

Select file to upload.
Files must be less than 2048 MB.
Allowed file types:

☐ Add text file to this upload?

Previous

Ingest