Book Content Model

Using the Book Content Model

Ingesting a book is a two step process. The first step involves creating a container for the pages of the book by ingesting the "Islandora Internet Archive Book Content Model", and the second step involves ingesting book page images into the Internet Archive Book object. There are numerous ways to ingest books into the repository, but CTDA recommends the following process as the easiest.

⚠️ In order to ingest books, you need to enable the Content Models in the collection policy of the collection object where the book will reside:
- Islandora Internet Archive Book Content Model
- Islandora Page Content Model

See the Choosing Content Models Guide for information about enabling Content Models.

Sections in this Guide

- Setting up the Internet Archive (IA) Book Viewer
- Adding Pages
- Adding Pages Using Zip Files
- Adding Multiple Zip Files to a Book
- Adding Books Using the Spreadsheet Ingest Module
  - Spreadsheet Specifications for Book Pages
- Managing Your Book
  - Reordering Pages
- Video Guide: Book Content Model
- Datastreams for Internet Archives Book Objects
- Datastreams for Book Page Objects
- Examples

Setting up the Internet Archive (IA) Book Viewer

- Log in to manage.ctdigitalarchive.org.
- Navigate to your collection on the CTDA Manage site.
• Click the "Manage" tab.

• Click "Add an object to this collection".

• Select "Islandora Internet Archive Book Content Model" from the "Select a Content Model to Ingest" dropdown menu and click "Next".

If you do not see the content model you are looking for, it may not be a part of your collection policy. You can add a content model to your collection policy by following the steps in this guide.
Select the MODS form you wish to use.

Choose your MODS form

Please select the correct MODS form you wish to use, as the different forms have defaults built into them. Most CTDA users, including users managed by Bibliomation, should use the CTDA Collection MODS Form or the CTDA Item MODS form. Users from the University of Connecticut and the Connecticut State Library should select the appropriate MODS form for their object.

You will be taken to the CTDA MODS form.

CTDA MODS Form

The CTDA MODS Form only has five required fields. But you should try to be thorough as possible and fill out as many fields as you can when describing your object. To learn more about the CTDA MODS Form, see the Describing Your Objects Guide.
- The first required field you need to populate is the title of your book.

![Title Information (Required)](image)

- The next required field is the "Resource Type", which should be "text".

![Resource Type (Required)](image)

- The third required field in the form is "Digital Origin". You should leave "reformatted digital" in this field.

![Format Description (Required)](image)

- The fourth required field is the "Held By" field. Select your institution from the dropdown menu.

![Held By (Required)](image)

- The final required field is the "Rights" field. Choose your institution from the dropdown menu. If you do not have a customized rights statement, select "Creative Commons NC-BY".
Click “Next” and the next page is where you will upload your object.

- You now have the option to upload PDF pages, but we recommend that you do not upload your pages at this point (see Adding Page procedures below).

- Scroll down to the bottom of the page select the language of the book and click “Ingest”.

**Rights Statements**

The default rights statement is set to the Creative Commons NC-BY 4.0 which states: “This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, CC BY-NC.” To learn more about Creative Commons licenses, visit [https://creativecommons.org/](https://creativecommons.org/).

If you have a custom rights statement, please send the text to us at ctda@uconn.edu.
Once ingest is completed, you will be taken to the Book viewer. From the Book viewer, you can add pages into the book and manage the book object.

Adding Pages

Once your Book Object has been ingested click the "Manage" tab on the page containing the book viewer.

Click "Book".

- Add Page
- Add Zipped Pages
• Select "Add Page".

• Select the MODS form you wish to use.

• Fill out the MODS form and click "Next".

• Click "Choose file" and find the page image file you want to add to the book.
Once you have chosen your file, click "Upload".

Finally choose the language of the text in the page image click "Ingest".

After your page image file has been ingested, you will be taken to the page displaying your file.

Adding Pages Using Zip Files
- Once your Book Object has been ingested click the "Manage" tab on the page containing the book viewer.

- Click "Book".

- **Add Page**
  - **Add Zipped Pages**

- **Add Page**
  - **Add Zipped Pages**

- Click “Add Zipped Pages”.

- **Add Page**
  - **Add Zipped Pages**

- Click “Choose File” and find the zip file you wish to add. Click "Upload".
• Choose the language of the book from the dropdown menu.

• Click “Add files”.

• A status bar will appear while your pages are added to the book. This status bar shows the progress of each individual page file being added to the book.

File Formats
The folder must be a .zip file to be uploaded. The pages contained in the .zip file can be .jp2, .jpeg, or .tiff files.

Click here to learn how to create zip files on a Windows computer.
Click here to learn how to create zip files on a Mac computer.

Upload and Ingest Time
Depending on the file size, the upload can take a few moments to a few minutes. Do not click “Back” or “Refresh” while the ingest is taking place.

Depending on the size and number of files contained in the zip file, the ingest process can take a few minutes to a half an hour in some cases. Do not click “Back” or “Refresh” while the ingest is taking place.
Adding Multiple Zip Files to a Book

Since the maximum file size of a zip file that can be uploaded to a book is 2GB, you may need to upload large books in multiple zip files.

- After you upload your first zip file (see above) to your book, return to the book viewer and click the “Manage” button.

- Select “Book”.

- Add Page
- Add Zipped Pages

- Click “Add Zipped Pages”.

- Add Page
- Add Zipped Pages
• Choose the language of the book from the "Language" dropdown menu.

• The "Last Sequence Number" will display the number of pages currently in the book. We recommend that you do not change the number in the "Last Sequence Number" field.

• Next click "Choose file" and locate the zip file you want to add.

• Once you have found your file, click "Upload".

! Changing the "Last Sequence Number"
Changing the last sequence number will result in the new individual page files being placed in between individual pages already in the book causing the pages to appear out of order. If you have uploaded a large book out of sequence, it is easier to keep uploading the zip files without changing the "Last Sequence Number" and manually reorder the pages after all of the files have been added to the book. To learn how to reorder book pages, see the Managing a Book Guide.
Adding Books Using the Spreadsheet Ingest Module

Users can now add newspapers to the repository by using the spreadsheet ingest module.
Please review the steps you need to take to connect to the CTDA network and servers prior to using the module.

Book objects can be added using the Standard Template or the Linked Data Template.

**Spreadsheet Specifications for Book Pages**

In the "isPageOf" and the "parent_object" columns, enter the PID of the book to which you are adding the book pages.

In the "parent_predicate" column, select "isMemberOf" from the dropdown list.

In the "isPageNumber" and "isSequenceNumber" columns, enter the page number of each page image in the order you want them to appear in the book.

In the "cmodel" column, select "islandora:pageCModel" from the dropdown list.

Book objects and pages cannot be added using the same spreadsheet. Users need to ingest the book object before adding pages to the book.

After you have filled out the appropriate columns in your spreadsheet, return to the Ingest Workflow guide to continue the spreadsheet ingest process.

**Managing Your Book**

**Reordering Pages**

- Log in to manage.ctdigitalarchive.org.
- Navigate to your book and click the "Manage" tab.
- Click "Book".
  - Add Page
  - Add Zipped Pages
- Select “Reorder Pages”.

<table>
<thead>
<tr>
<th>Page #</th>
<th>Thumbnail</th>
<th>Label</th>
<th>Page #</th>
<th>Thumbnail</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>index-1_1</td>
<td>1</td>
<td></td>
<td>index-1_1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>index-2_1</td>
<td>2</td>
<td></td>
<td>index-2_1</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>index-3_1</td>
<td>3</td>
<td></td>
<td>index-3_1</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>index-4_1</td>
<td>4</td>
<td></td>
<td>index-4_1</td>
</tr>
</tbody>
</table>

- The pages will appear in two columns, side-by-side, and display 10 pages per column by default. You can drop and drag pages in the same column and in between columns.
- Scroll down to the bottom of the columns and you will be able to view more pages by clicking on the number list at the bottom of the columns.

- The pages you move will show yellow check marks next to the page image.

- Once you have reordered the pages in the correct order, scroll down and click “Save Changes”.

**Quick Reorder**

**Move Pages:**

Must follow the form ‘1, 2-4,5’.

**Insert After:**

Go

Save Changes
We recommend that you do not use the "QUICK REORDER" options. The process outlined above is more straightforward.

- A status bar will appear while your pages are added to the book.

Depending on the number of pages you have reordered, the process can take a few moments to a few minutes. Do not click "Back" or "Refresh" while the reorder is taking place.

- When completed, you will receive a "Sequenced Book" confirmation message.

Video Guide: Book Content Model

Coming Soon!

Datastreams for Internet Archives Book Objects

<table>
<thead>
<tr>
<th>ID</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELS-EXT</td>
<td>Default Fedora relationship metadata</td>
</tr>
<tr>
<td>MODS</td>
<td>MODS metadata form</td>
</tr>
<tr>
<td>DC</td>
<td>Dublin Core record</td>
</tr>
<tr>
<td>TN</td>
<td>Thumbnail image</td>
</tr>
<tr>
<td>PDF</td>
<td>PDF derivative created by ImageMagick</td>
</tr>
</tbody>
</table>

Datastreams for Book Page Objects

<table>
<thead>
<tr>
<th>ID</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBJ</td>
<td>Original TIFF or JP2 file uploaded</td>
</tr>
<tr>
<td>DC</td>
<td>Dublin Core record</td>
</tr>
<tr>
<td>PDF</td>
<td>PDF derivative created by Ghostscript</td>
</tr>
<tr>
<td>JP2</td>
<td>JPEG 2000 derivative</td>
</tr>
<tr>
<td>JPG</td>
<td>Smaller JPEG derivative</td>
</tr>
<tr>
<td>TN</td>
<td>Thumbnail icon created from the image during the ingest process</td>
</tr>
<tr>
<td>RELS-INT</td>
<td>Internal Fedora relationship metadata defining the dimensions of the JP2 datastream</td>
</tr>
<tr>
<td><strong>OCR</strong></td>
<td>Optical Character Recognition file</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td><strong>HOOCR</strong></td>
<td>A converted version of the OCR datastream, intended to be more human-readable</td>
</tr>
<tr>
<td><strong>RELS-EXT</strong></td>
<td>Default Fedora relationship metadata</td>
</tr>
</tbody>
</table>

**Examples**