Books Solution Pack

Overview

The Book Solution pack creates a Book Collection object consisting of a MODS metadata records. After the Book Collection object is created, users can upload a zipped directory of uncompressed tiffs. These tiffs become Page objects that are members of the Book Collection Object. Page objects undergo an Optical Character Recognition (OCR) process, making their contents full-text searchable. You can use the Book module to create any paged content consisting of tiff images of pages.


While books for Islandora 6 and 7 can exist in the same repository, books ingested through Islandora 6 will not display properly in Islandora 7. A migration script is required.

Tutorials

Creating and Adding Pages to a Book

Managing Book Objects

Managing Page Objects

Migrating Books from Older Versions of Islandora

Dependencies

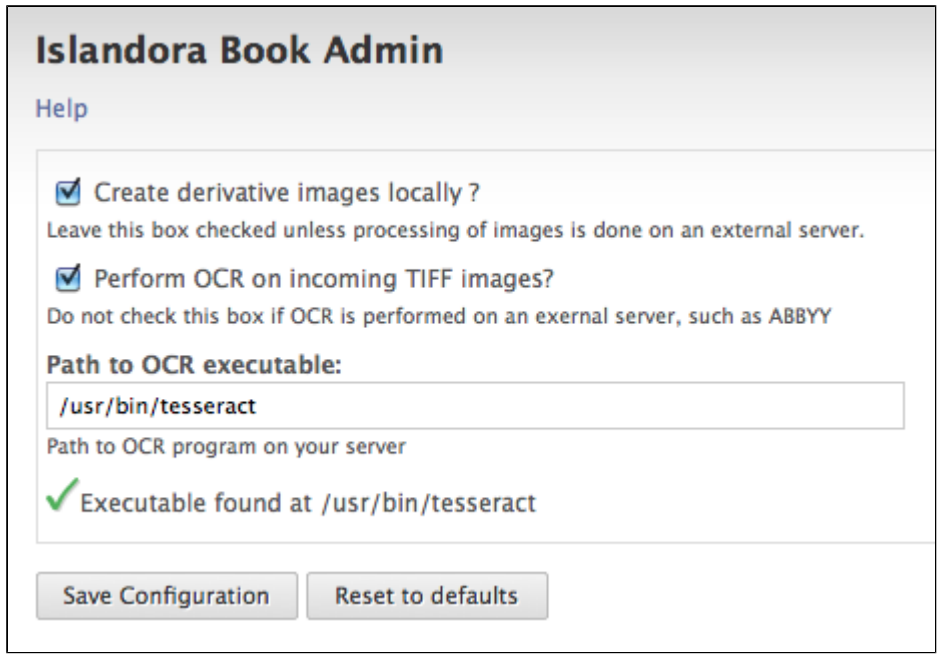
- If you are performing OCR services locally, you will need to install Tesseract 3.0.
- Kakadu / Djatoka
- ImageMagick
- XML Forms and dependencies for that module
- Islandora Image Viewer

Configuring the Book Solution Pack

After you have activated and installed the Book Solution Pack, you may choose whether or not to use local services to create page image derivatives (Kakadu/Djatoka) and whether or not to use local services to perform Optical Character Recognition (OCR) on your ingested, uncompressed .tiff images.

If you wish to use microservices to perform OCR services using an external service, you will want to deselect these features.

Once you have activated and installed the book solution pack, you can navigate to administer> Islandora Book Admin (under “Site Configuration). You will be presented with the following screen:



If you wish to use external services for image derivation and OCR, deselect these boxes. You will need to make sure that the path to your version of tesseract (that has been installed in the same server) is identified in "Path to OCR executable." A green check mark will appear to indicate that the system has discovered the tesseract directory.

Content Models and Prescribed Datastreams

The Books Solution Pack uses the following content models:

- islandora:pageCModel
- islandora:bookCModel
- islandora:jp2Sdef

The Book Object will be created with the following datastreams:

RELS-EXT	Default Fedora Relationship Metadata Stream
DC	Default Dublin Core Data Stream
TN	Thumbnail for Display
MODS	Data Stream holding MODS Metadata
PDF	PDF copy of the book

** Note that there are no binary files in the Book Object, because the book object only stores the bibliographic information for the book.*

Page Objects will be created with the following datastreams:

RELS-EXT	Default Fedora Relationship Metadata Stream
DC	Default Dublin Core Data Stream
TN	Thumbnail for Display
JP2	Web-viewable JPEG2000 image
TIFF	Original Binary file
OCR	OCR text stream

The default form is called “IslandoraBooksMODSForm”

Metadata Field Mappings

The metadata fields displayed on the 'Description' tab of a book object are populated from MODS fields based on the specifications in **book_view.xsl**. The mappings are as follows:

Description Field	Form Field	MODS Element
By Statement	Statement of Responsibility	//mods:note
Place of Publication	Origin Information > Place	//mods:placeTerm
Publisher	Origin Information > Publisher	//mods:publisher
Date	Origin Information > Date Issued	//mods:dateIssued
Language	Language	//mods:languageTerm
Pagination	Physical Description > Extent	//mods:extent
ISBN 10	ISBN	//mods:identifier[@type='isbn']
Subjects	Subject	//mods:subject