

MULTIPAGE OBJECTS

VITA 6.4

Introduction	3
Creating Multipage Digital Objects	3
Uploading Multipage Objects	3
Adding A "Hero" Shot	3
Creating Records For Text Objects	4
Adding Pages From Multiple Files	5
Step 1: File Names	5
Step 2: Uploading Multiple Files	6
Adding Pages From A Single PDF	9
Step 1: Add Record With File	9
Step 2: Splitting Pdfs	11
Managing Multipage Pages & Files	12
Batch Relabelling Of Pages	12
Reordering Pages	12
Batch Deleting Pages	15
Editing Page File Details	17
Replacing Individual Pages	19
Download PDF Options	20
Page-Level Downloads	20
Object-Level Downloads	20
Multipage Display	20

INTRODUCTION

This manual shows how to upload and manage multipage items like books, scrapbooks, brochures, pamphlets, etc. See the Adding and Managing records manual for guidance on uploading single-page documents; see the Indexing Articles & BMDs to see how to manage collections of text fragments and Newspaper Publication & Issues for full run newspaper collections.

VITA text management features help you upload and manage the variety of complex text objects you might find in your collections for better discovery and display. Use this manual to walk through the process of uploading multipage objects either from image files or PDFs.

CREATING MULTIPAGE DIGITAL OBJECTS

- Scan each page of your object separately
- Scan text pages at 300dpi and save in TIFF format for master copies
- Name the scans for each object with a consistent naming convention e.g. BookTitle_oooo1.tif, BookTitle_oooo2.tif, BookTitle_oooo3.tif... etc.
- See the File Naming Conventions document for more information
- Save the TIFs again as JPGs, JPEG2000s(JP2)or PDFs
- Upload your JPG, JP2 or PDF to VITA for full text search and hit highlighted results
- NOTE: .doc, .rtf and .txt files will not render hit highlighted results unless PDF-d first.

UPLOADING MULTIPAGE OBJECTS

Like all records, you want to establish the record with a representative image for the object or a "hero shot". If you are uploading a multipage PDF, the cover page will automatically be assigned this role (and can be overwritten – see below). For multipage items that consist of a batch of single page files, add a "hero shot" first, then load the pages.

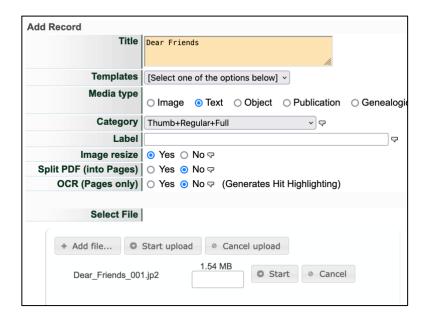
ADDING A "HERO" SHOT

If adding a PDF that will be Split, you can skip this step so long as the first page is an appropriate "hero" image for the record (or you can overwrite the thumb and regular image for the record after the fact).

Otherwise, create a display image for the multipage object:

Go to the Main Menu→Add record with file

- Add Title
- Media Type: "Text" assuming this describes the primary content of the multipage object (You can add Secondary Media Types later)
- Category: Select "Full+Reg+Thumb" or "Reg+Thumb"
- Label: Not applicable
- File re-size: Select default "Yes"
- Split PDF: keep default No for hero shots
- Browse for the file you want to represent this multipage object (e.g. a front cover or first page)



- Click Start Upload
- When file is loaded, click the "Continue" button

CREATING RECORDS FOR TEXT OBJECTS

As soon as an image is uploaded, you can start to create the metadata for the text object.

For full instructions on adding metadata and using the VITA templates, see the Adding & Managing Records or the VITA Metadata Application Profile documents.

ADDING PAGES FROM MULTIPLE FILES

Go to the File/Tech Data screen. You will see your initial image listed like this



To add pages to your record, click on "Associate multiple page files with this record" and follow these steps:



STEP 1: FILE NAMES

Here you can save a lot of time by automatically assigning page labels by "stripping" the common or "consistent" parts of the file names for all the page files you are uploading. Do not include preceding zeros from the file name as these will be stripped out automatically.

Add pages, step 1		
Title	Dear Friends	
File Name constants	Dear_Friends_ Prefix .jp2	Suffix
	If your pages have a consistent namin "Book") and suffix (e.g. ".tif") in the boo to Label the page. You will be able to f	kes above. Th
Change file size (page images only)	Yes ○ No	
OCR / Hit highlighting	Yes ○ No ▽	
	on to Step 2	

E.g. the file names for my pages files **all** include "Dear_Friends_" (the prefix) and zeros that precede the page number; all the files are JP2s, so the extension or suffix is ".jp2".

Note: we recommend uploading page files as JPG 2000s because of the improved panzoom and rotate options. See the Working with JP2 manual for more information.

Allow the tool to resize your image files (settings for page size can be reviewed in Agency Management>Image Options)

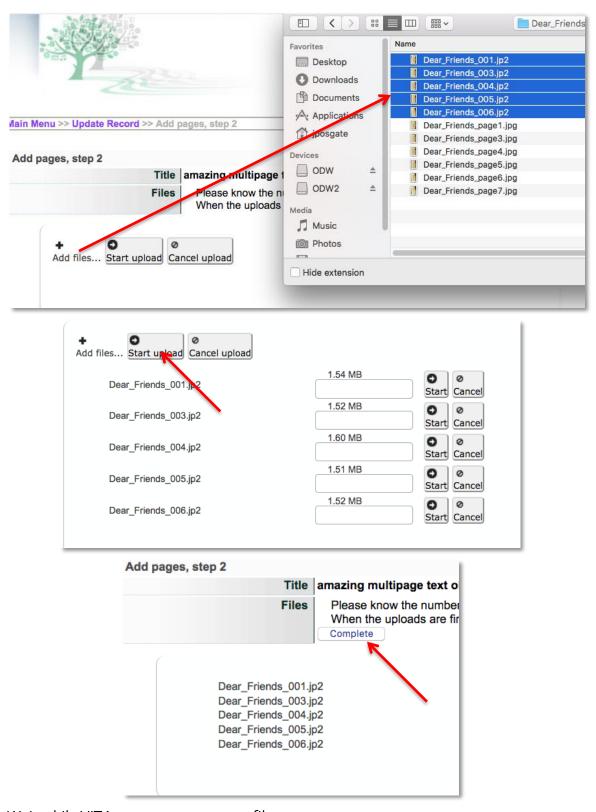
OCR /Hit highlighting – choose "yes" if your pages have typed content. The tool will generate a text file as well as apply hit highlighting for the words in each page automatically. For handwritten material, choose "no".

Click on "...on to Step 2" button.

STEP 2: UPLOADING MULTIPLE FILES

This is a batch file upload screen. Choose the files from your local drive and click "Start uploading" then click the "Continue" button once all the files have completed their load.

Add Record		
Title	amazing <u>multipage</u> item	
Templates	[Select one of the options below] ~	
Media type	○ Image	○ Genealogi
Category	Page (Multiple allowed)	
Label		₽
	Yes ○ No ▽	
Split PDF (into Pages)	O Yes ⊙ No ⊽	
OCR (Pages only)	Yes ○ No ♥ (Generates Hit Highlighting)	



Wait while VITA processes your page files:

Note: It takes approximately 1 minute per page to generate all the display images as well as OCR'ing the page content and applying to coordinates that will produce the hit highlighting on the page text when a search is performed. As well, it will take up to 2 ominutes for the full text and hit highlighting to be active on your site depending on when the new material enters the indexing cycle.

Jess Posgate (Test

Don't navigate away from this page until the "continue" message appea



Ottewa, Feb. The work of the Bee tot is mov TM Sa to 10003 as thongs we H Government Mavisy. The Supplementary Estimates of went t rough ina day, '



oe think we must 911 sdmit, I think even the est pertisen, even the mcst blind ensible government once meont being respensible to the crown. Hewever, gr

Then click "Continue"

6 processed

You completed Step 3 by adding 6 files at 17 Aug 2022 16:13:18 Continue...

New files will display in the File/tech data screen and can then be edited individually.



Note that the page labels match the unique variables in the file names, but with the full file name stripped away.

ADDING PAGES FROM A SINGLE PDF

Adding a single PDF for an issue means you do not have to strip the file names and the text is automatically extracted for full text search with hit highlighting.

STEP 1: ADD RECORD WITH FILE

- Go to the Main Menu→Add record with file
- Add Title
- Select Media Type "Text" assuming this describes the primary content of the multipage object (You can add Secondary Media Types later)
- Select "Page" from the Category drop down list
- Adding a Label for the full object download (Optional)
- File re-size: Select default "Yes"
- Split PDF: "Yes"
- OCR: This will happen automatically for pages generated from split PDFs
- Browse for the file you want to upload



Wait for file to load, then click "Continue"

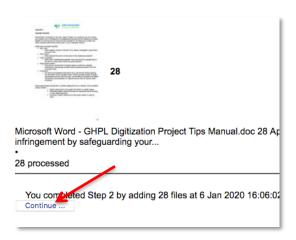
STEP 2: SPLITTING PDFS

On the next screen, the tool needs to process the single PDF into multiple pages, extract the text and create the Thumbnails and other associated files for each page.

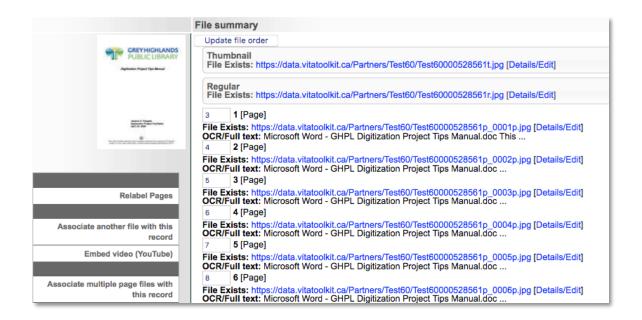
Note: It takes approximately 1 minute per page to generate all the display images as well as OCR'ing the page content and applying to coordinates that will produce the hit highlighting on the page text when a search is performed. As well, it will take up to 20 minutes for the full text and hit highlighting to be active on your site depending on when the new material enters the indexing cycle.



Wait for all the files to be processed completely and then click the button at the bottom of the screen:



On the File/Tech screen, you'll see the files generated from your PDF, including the Thumbnail and Regular image files, every page with extracted text snippets, and a complete PDF that is automatically associated with the record. You can make the full PDF public or not, depending on whether you want to allow downloading.



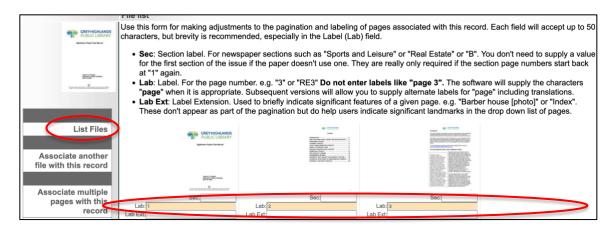
MANAGING MULTIPAGE PAGES & FILES

BATCH RELABELLING OF PAGES

Click the Relabel Pages link

Add new label numbers for all pages.

Optional (for newspaper issues): add section name (e.g. Seasonal Insert, Sports, etc)

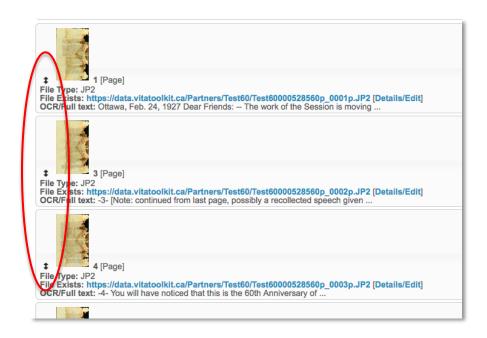


Click **Update Labels** to commit changes

To return to File/Tech data screen, click List Files link in sidebar

REORDERING PAGES

For issues with less than 20 pages, use the elevator buttons by clicking and dragging to reorder pages as required.



For items with more than 20 pages, the page "order" will appear beside the page label in a numeral box. This order starts with the item Thumbnail (assigned order 1), then the Regular (assigned order 2), and then Page 1 (assigned order 3).

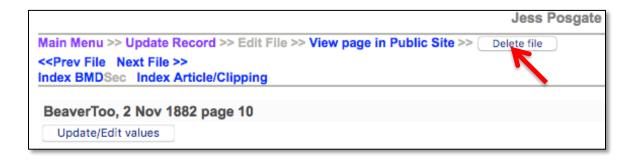
To reorder, either renumber the whole set and click update or, to insert one page between two others, use ##.5, i.e. if the first page is somehow buried in the rest of the pages, assign the page order number 2.5 and click update. The first page will then be moved into the #3 position and the rest of the pages will follow; unless the labels were correct for each page file, you will need to relabel the pages if the order is changed. See Group management for more details.

```
1 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0001p.JP2 [Details/Edit]
OCR Full text: P, "BoOoe|1 LECescTT Loaws a the we*#t;f'Te';§s...
          2 [Page] Non public
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0002p.JP2 [Details/Edit] OCR/Full text: (yg'ntjiii'h,','iai..ii,rti I)yriaE',l'IIIIIPrII.""" III. f . 'I'A2tti ...
          3 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0003p.JP2 [Details/Edit]
OCR/ffull text: i, . - - - ,...' (5'73"! 'yr new: f .1 cJ..- A at?» , i2 - ...
          4 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0004p.JP2 [Details/Edit]
OCR/Full text: n. member for Frontsuac. Hitherto the 8ir AENRY SMITH Government _ ...
          5 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0005p.JP2 [Details/Edit]
OCR/full text: --" ' yii-T V J," if"? , -, ' . debate . we all members returned upon ...
          6 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0006p.JP2 [Details/Edit]
OCR/Full text: SE T U PE hx Ho y q » wihie )ns Hon, J. S, MNcDONALD intimated that ...
          7 [Page]
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0007p.JP2 [Details/Edit]
OCR/Full text: NOR SUBMIT TO ARBTTRARY WEASUIET nE TT e ie atermetstrmmeepe e ...
 10
          8 [Page]
File Type: JP2
Rile Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p 0008p.JP2 [Details/Edit]
OCR/Full text: irom suon old mising district, compared with th»] 3 xpense of collecting ...
          9 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0009p.JP2 [Details/Edit]
OCR/Full text: .";y 'te bey on d #, Lias 1t devoived on every memb--r of tha Hovse to ...
           10 [Page]
File Type: IP2
```

BATCH DELETING PAGES

For issues with less than 20 pages, delete the individual pages by opening the details/edit link and using the delete file button.





For items with more than 20 pages, the page "order" will appear beside the page label in a numeral box. This order starts with the issue Thumbnail (assigned order 1), then the Regular (assigned order 2), and Page 1 (assigned order 3).

To batch delete pages, enter "D" in the numeral box for the files you would like to delete and click Update. This is an irreversible decision, so be certain you want those files removed from the record and the server.

```
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528561p_0025p.jpg [Details/Edit]
OCR/Full text: Microsoft Word - GHPL Digitization Project Tips Manual.doc ...

D 26 [Page]
File tx sts: https://data.vitatoolkit.ca/Partners/Test60/Test60000528561p_0026p.jpg [Details/Edit]
OCR/Full text: Microsoft Word - GHPL Digitization Project Tips Manual.doc ...

D 27 [Page]
File tx sts: https://data.vitatoolkit.ca/Partners/Test60/Test60000528561p_0027p.jpg [Details/Edit]
OCR/Full text: Microsoft Word - GHPL Digitization Project Tips Manual.doc ...

D 28 [Page]
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528561p_0028p.jpg [Details/Edit]
OCR/Full text: Microsoft Word - GHPL Digitization Project Tips Manual.doc ...

31 download this whole thing [Text file] Non public
File Type: PDF
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test600005285611T.PDF [Details/Edit]
OCR/Full text: Microsoft Word - GHPL Digitization Project Tips Manual.doc This ...

Update file order
```

EDITING PAGE FILE DETAILS

On the File Details/Edit screen you can edit labels, settings, and add or replace files.



Section: Use to denote a specific section, such as a Chapter heading or Volume number. This will precede the page number in the page drop down list.

Label: Any label information will be carried forward from the Add Record screen. Edit Labels here.

Label (extension): Use this to distinguish a labelled page as something else, like Title Page, or Index. The file display will include a colon and the extension label (e.g. 1: Title Page).

Category: The category displayed here reflects either the category selected on the Add Record screen (e.g. Text, Detail, etc) or will be automatically populated during a batch upload of pages (i.e. Category: Page).

Files: Shows all the files created for and associated with that Page. The original digital file name, an option to add or replace a thumbnail and other display files for the record and/or sidebar display, the file name and file Replace option, and File Size in bytes or pixels, and File type. (See below on replacing files)



Put in Public Display? Default is "Yes". Use this option to manage files that are not meant for public display such as specifically sized details files, or the text files associated with an object where you do not want the text file to be displayed but want it's content to be indexed for full text discovery.

Index the text in this file? This allows you to select which of multiple files will have their content indexed for full text discovery. E.g. if you upload a PDF object, the text will automatically be extracted, but if you upload a DOC file with the corrected text, then select "No" for the PDF and "Yes" for the DOC, and only the DOC file content will be indexed.

Full Text: Any full text that is auto-extracted from the file you uploaded will appear in this input box; use this input box to manage, edit or review full text. This full text will not display in the public details of the object. **Note:** Altering this text will not alter the hit highlighting on search results, only full text page-level "hits"

Remember: Click "**Update/Edit Values**" on this screen to commit your changes on this page.

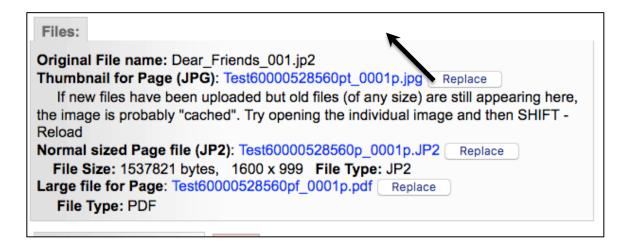
REPLACING INDIVIDUAL PAGES

If you want to replace a page or any of the derivative display files for that page, go to the file/tech data screen and click the details/edit link on the page in question.

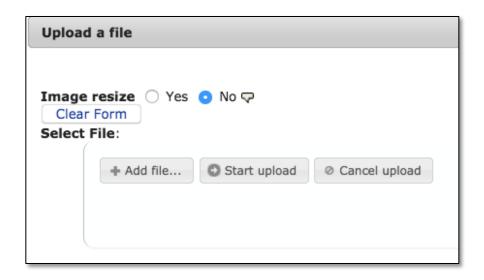


In the Files field, you can upload a new Thumbnail, Page image file, and the PDF file for that page (these PDFs are automatically derived from your original page file and have no full text).

Replacing the page file will not disturb any links to index records like articles or BMDs. If you do not see your new page, be sure to clear your browser cache and/or do a hard refresh.



NOTE: When replacing pages with JP2 files, choose "No" to the Image resize option



DOWNLOAD PDF OPTIONS

PAGE-LEVEL DOWNLOADS

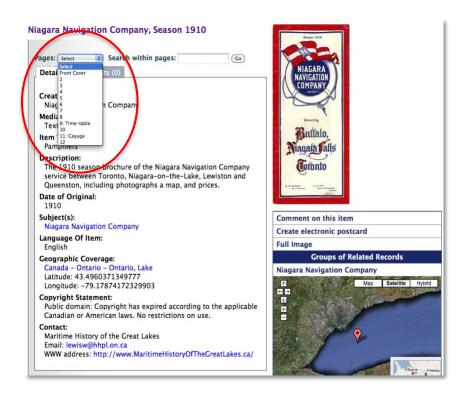
Single, readable PDF files are automatically generated for every "Page" in VITA and offered for page-level download for the end user. If you would like to change that option, see the Agency Management manual.

OBJECT-LEVEL DOWNLOADS

When a single PDF is uploaded and Split, the original PDF is preserved and added to the record. The default display setting for complete PDF objects is "not public." If you want to offer the whole item for download, go to File/Tech data screen, select the details/edit for the single PDF file and change the status of that file. The download option displays in the record's right sidebar; you can also change the label of that file to something non-generic.

MULTIPAGE DISPLAY

When multiple files are associated with one record, the public display of the internal pages will render like this:



Page viewer will look like this, with a search bar for full-text search:

