Part 1	**Text Object Features**

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, PDFs, full text extraction, OCR processing, and Zoomify display. Also, see the section on Clippings and Publications to see how to manage collections of text fragments.

Part 2	**Multipage Objects**

Creating the digital object:

- Scan each page of your object separately
- Scan text object at 300dpi and save in TIFF format for master copies, and derive JPGs for OCR processing, and upload JPG and/or JP2000 to VITA
- Name the scans for each object with a consistent naming convention e.g. BookTitle_00001.tif, BookTitle_00002.tif, BookTitle_00003.tif…etc.
- See the File Naming Conventions document for more information

Part 2.1	**Display image for multipage objects**

First, create a display image for the multipage object:

Go to the Main Menu→Add record with file
- Include Title
- Browse for the file you want to represent this multipage object (e.g. a front cover or first page)
- Select Media Type "Text" or "Image" depending on the content of the multipage object
- Select "Full+Reg+Thumb" or "Reg+Thumb" from the category drop down list
- Label is unnecessary for this category
- File re-size: Select default "Yes"
- Split PDF: default is No unless you are working with a multipage single file PDF (See Section 3)
- Click upload
- On the File/Tech Data screen the three display files will be listed like this:

Part 2.2   Uploading multiple pages

Next, add the multiple pages of the object:

- On the File / Tech Data screen, click on "Associate multiple pages with this record" and follow these steps:

Step 1: File Names

Input the consistent prefix (e.g. BookTitle_ID_) and suffix (e.g. .jpg) that can be associated with multiple files.
Remember: include any trailing punctuation in the prefix and the period in the suffix.

Click on "…on to Step 2" button.

For example: Your file names are BookTitle_ID_00001.jpg, BookTitle_ID_00002.jpg, BookTitle_00003.jpg...etc. with each page identified by the 000#. To reduce the label to the page number only (which will display in the dropdown list in the display), enter "BookTitle_ID_" into the prefix box and ".jpg" as the suffix. The resulting display will be 1,2,3, etc. This is what will show in the drop down menu on the public display side (see Section 2.5: Multipage Object Display)

Step 2: Browse and Select Files

Click Add files

Choose files for uploading from your computer (highlight multiple files for uploading using the Command+Click (Mac) or Ctrl+Click (PC) keyboard commands) then use the "Start Upload" button at the top of the screen to do a batch upload.

Title	Dear Friends
Files	Please know the number and size of files you were attempting to upload when reporting problems. When the uploads are finished click the "Complete" button to advance to Step 3.

+ Add files...   ⊙ Start upload   ⊘ Cancel upload

	345.59 KB	⊙ Start	⊘ Cancel
Dear_Friends_page1.jpg			
Dear_Friends_page3.jpg	341.45 KB	⊙ Start	⊘ Cancel
Dear_Friends_page4.jpg	375.96 KB	⊙ Start	⊘ Cancel
Dear_Friends_page5.jpg	334.81 KB	⊙ Start	⊘ Cancel
Dear_Friends_page6.jpg	340.65 KB	⊙ Start	⊘ Cancel
Dear_Friends_page7.jpg	241.47 KB	⊙ Start	⊘ Cancel

The selected files will show progress bars.

Title	Dear Friends
Files	Please know the number and size of files you were attempting to upload when reporting problems. When the uploads are finished click the "Complete" button to advance to Step 3.

+ Add files...   ⊙ Start upload   ⊘ Cancel upload

150.67 kbit/s | 00:00:51 | 50.94 % | 1.01 MB / 1.98 MB

	345.59 KB	⊙ Start	⊘ Cancel
Dear_Friends_page1.jpg			
Dear_Friends_page3.jpg	341.45 KB	⊙ Start	⊘ Cancel
Dear_Friends_page4.jpg	375.96 KB	⊙ Start	⊘ Cancel
Dear_Friends_page5.jpg	334.81 KB	⊙ Start	⊘ Cancel
Dear_Friends_page6.jpg	340.65 KB	⊙ Start	⊘ Cancel
Dear_Friends_page7.jpg	241.47 KB	⊙ Start	⊘ Cancel

When all progress bars reach 100%, click the "Complete" button.

Add pages, step 2

Title	Dear Friends
Files	Please know the number and size of files you were attempting to upload when reporting problems. When the uploads are finished click the "Complete" button to advance to Step 3.

Complete

Dear_Friends_page1.jpg	345.59 KB
Dear_Friends_page3.jpg	341.45 KB
Dear_Friends_page4.jpg	375.96 KB
Dear_Friends_page5.jpg	334.81 KB
Dear_Friends_page6.jpg	340.65 KB
Dear_Friends_page7.jpg	241.47 KB

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

File summary

Relabel Pages

Associate another file with this record

Associate a single Zoomify zip file with this record

Associate multiple pages with this record

Associate OCR/Full text files with these pages

Associate Zoomify files with these pages

When finished editing pages/text:

Update Index

Thumbnail
File Exists: https://data.ourontario.ca/Partners/Training/OOTr001638985t.jpg [Details/Edit]
Regular
File Exists: https://data.ourontario.ca/Partners/Training/OOTr001638985.jpg [Details/Edit]
Full Image
File Exists: https://data.ourontario.ca/Partners/Training/OOTr001638985f.jpg [Details/Edit]
1 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/OOTr0016389850001p.jpg [Details/Edit]
OCR/Full text: ...
Zoomify: files available
3 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/OOTr0016389850002p.jpg [Details/Edit]
OCR/Full text: ...
Zoomify: files available
4 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/OOTr0016389850003p.jpg [Details/Edit]
OCR/Full text: ...
Zoomify: files available
5 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/OOTr0016389850004p.jpg [Details/Edit]
OCR/Full text: ...
Zoomify: files available
6 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/OOTr0016389850005p.jpg [Details/Edit]
OCR/Full text: ...
Zoomify: files available
7 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/OOTr0016389850006p.jpg [Details/Edit]
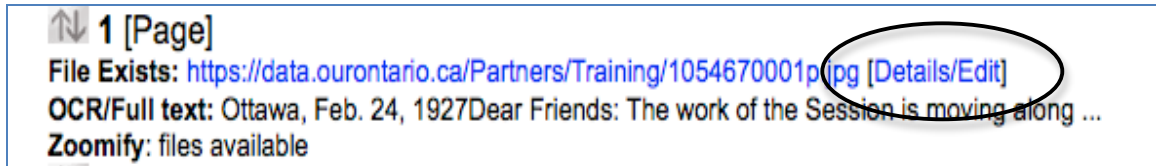**OCR/Full text:** ...
Zoomify: files available

To order the links click and drag the thumbnail or sort icon up or down the list.

If working with JPG or JP2/TXT files (as opposed to readable PDFs), repeat the above steps with the TXT files.

NOTE: the JPG/JP2 and TXT files must have the same file names to associate with one another.

Part 2.3	Manage page files

File details are automatically generated and can be reviewed by clicking on the Details/Edit link for each file on the File / Tech Data screen.
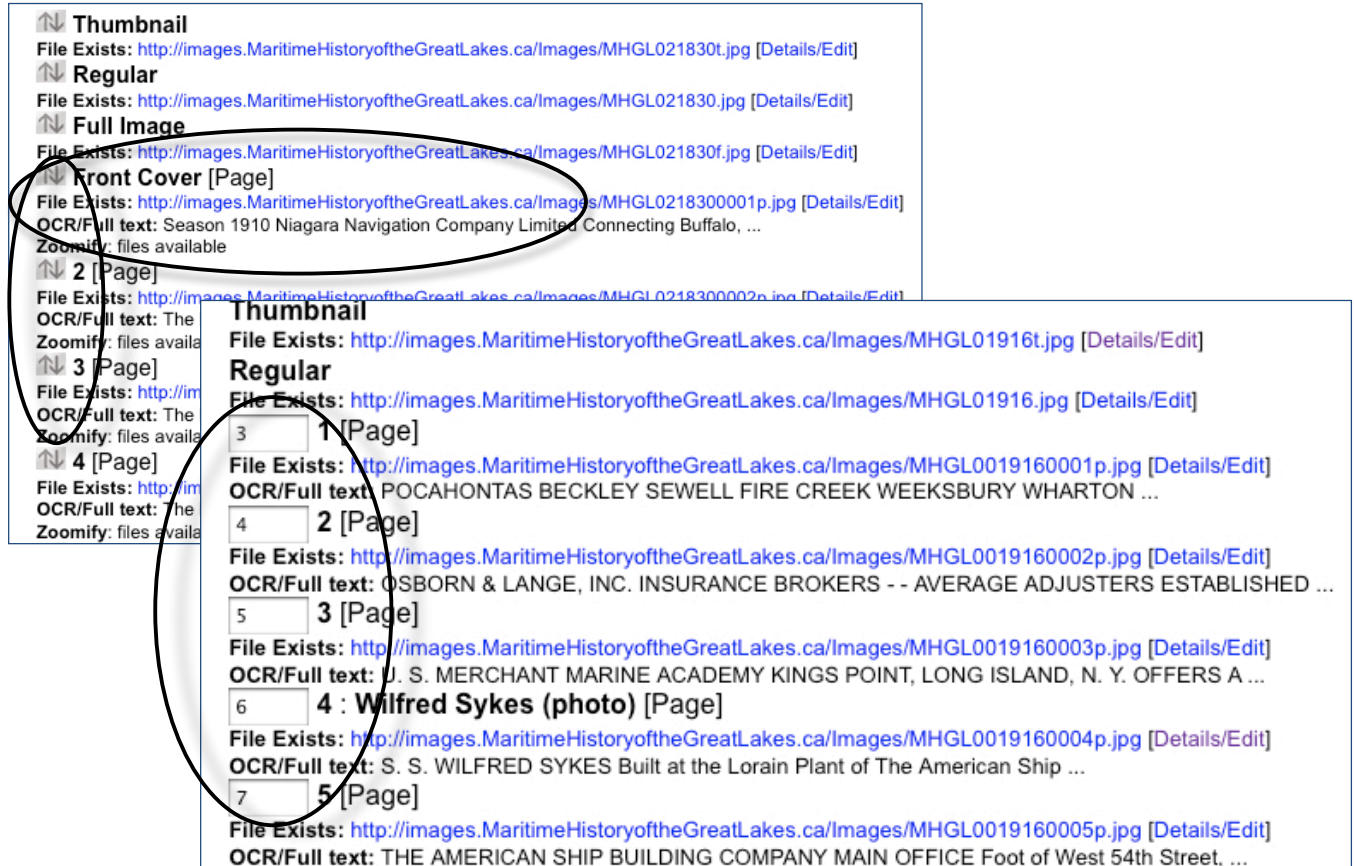
⇅ 1 [Page]
File Exists: https://data.ourontario.ca/Partners/Training/1054670001p.jpg [Details/Edit]
OCR/Full text: Ottawa, Feb. 24, 1927Dear Friends: The work of the Session is moving along ...
Zoomify: files available

- "File exists" provides the location of the digital object file.
- "OCR / Full text" shows a snippet of the text that is associated with the image file (if available)
- "Zoomify: Files available" indicates that there are Zoomify-generated files associated with the page or file listed here

File order
- ▪ **For objects up to 20 pages** order the links by clicking and dragging the "sort" icon up or down the list. This will commit your changes.
- ▪ **For objects with more than 20 pages** order the pages using the numeric input boxes. Click the Update File Order button to commit your changes.

⇅ **Thumbnail**
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL021830t.jpg [Details/Edit]
⇅ **Regular**
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL021830.jpg [Details/Edit]
⇅ **Full Image**
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL021830f.jpg [Details/Edit]
⇅ **Front Cover** [Page]
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0218300001p.jpg [Details/Edit]
OCR/Full text: Season 1910 Niagara Navigation Company Limited Connecting Buffalo, ...
Zoomify: files available
⇅ 2 [Page]
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0218300002p.jpg [Details/Edit]
OCR/Full text: The
Zoomify: files avail
⇅ 3 [Page]
File Exists: http://im
OCR/Full text: The
Zoomify: files avail
⇅ 4 [Page]
File Exists: http://im
OCR/Full text: The
Zoomify: files avail

Thumbnail
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL01916t.jpg [Details/Edit]
Regular
File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL01916.jpg [Details/Edit]

| 3 | 1 [Page] |

File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0019160001p.jpg [Details/Edit]
OCR/Full text: POCAHONTAS BECKLEY SEWELL FIRE CREEK WEEKSBURY WHARTON ...

| 4 | 2 [Page] |

File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0019160002p.jpg [Details/Edit]
OCR/Full text: OSBORN & LANGE, INC. INSURANCE BROKERS - - AVERAGE ADJUSTERS ESTABLISHED ...

| 5 | 3 [Page] |

File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0019160003p.jpg [Details/Edit]
OCR/Full text: U. S. MERCHANT MARINE ACADEMY KINGS POINT, LONG ISLAND, N. Y. OFFERS A ...

| 6 | 4 : Wilfred Sykes (photo) [Page] |

File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0019160004p.jpg [Details/Edit]
OCR/Full text: S. S. WILFRED SYKES Built at the Lorain Plant of The American Ship ...

| 7 | 5 [Page] |

File Exists: http://images.MaritimeHistoryoftheGreatLakes.ca/Images/MHGL0019160005p.jpg [Details/Edit]
OCR/Full text: THE AMERICAN SHIP BUILDING COMPANY MAIN OFFICE Foot of West 54th Street, ...

File "Edit"

To add or edit details associated with a file, click "Details/Edit" at the end of the File Exists URL

Dear Friends page 1

(Update/Edit values)

Label:

1

For the page number. e.g. "3" or "RE3" **Do not enter labels like "page 3".** The software will supply the characters **"page"** when it is appropriate. Subsequent versions will allow you to supply alternate labels for "page" including translations.

Label (extension):

Used to briefly indicate significant features of a given page. e.g. "Chapter 2" or "Barber house [photo]" or "Index". These don't appear as part of the pagination but do help users indicate significant landmarks in the drop down list of pages.

Category:

Page

Files:

Original File name: Dear_Friends_page1.jpg
Regular file (JPG): 1054670001p.jpg
  File Size: 165678 bytes, 1000 x 624 **File Type:** jpg

Zoomify:

Files have been uploaded.
Delete Zoom files

Put in public display?:

○ Yes ⊙ No

Is this corrected OCR?:

○ Yes ⊙ No

Index the text in this file?:

⊙ Yes ○ No

Turn this off if you have uploaded multiple versions of the same content (say a Word .DOC as well as a PDF file), so that we only index it once. **Note:** one of those files will need to be indexed.

Full text:

Ottawa, Feb. 24, 1927Dear Friends: The work of the Session is moving along rapidly, [sic] It begins to look as though we would be out of here at Easter or shortly after. The general

Alternate Sizes 400x600 540x900 800x900 1200x1200

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).

Original File Name: The name of the file uploaded to the tool. This is key information for mapping between the digital object online and the digital object you uploaded.

Regular File: The system file name for your uploaded file.

File size provides information on the file size in bytes.

File type provides information on the type of file – usually JPEG.

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.
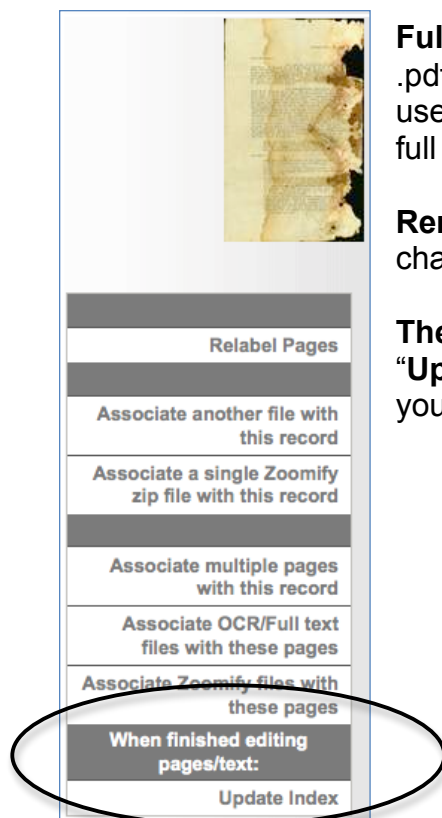
Is this corrected OCR?
If your full text is correct (i.e. edited for correct spelling, etc) , select "Yes". This allows the text from your page to be added to the index that drives the "Did you mean?" search help for end users.

Whether you select Yes or No, the full text displayed on this page is used for full text searching.

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 .doc, .pdf, .txt, .xsl, and .rtf objects will appear in this input box; use this input box to manage, edit or review full text. This full text will <u>not</u> display in the public details of the object.

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

Then: Return to the File/Tech Data screen and click the "**Update Index**" button to manually re-index the content you have uploaded or changed:

Part 2.5   Multipage display

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