

# MANAGING NEWSPAPER PUBLICATIONS & ISSUES

**VITA 6.1** 

UPDATED DECEMBER 2019

INTRODUCTION	
NEWSPAPER PUBLICATION RECORDS	4
ADD PUBLICATION RECORD	4
EDIT PUBLICATION TITLE	6
Publication record Management	6
ADD PUBLICATION MASTHEAD	6
ADDING ISSUES	7
STEP 1: ADDING ISSUES	8
STEP 1: UPLOADING MULTIPLE FILES	9
STEP 1: UPLOADING A SINGLE PDF	10
STEP 2: FILE NAMES	10
STEP 2: PAGE DISPLAY SIZE	11
STEP 3: MULTIPLE FILE UPLOADER	
STEP 4: MULTIPLE FILES PROCESSING & CONFIRMATION	
STEP 4: SINGLE PDFs PROCESSING & CONFIRMATION	14
MANAGING ISSUE FILES & CONTENT	16
EDIT ISSUE RECORD	16
ADDING FULL TEXT FOR MULTIPLE FILE PAGES	17
BATCH RELABELING OF ISSUE PAGES	
REORDERING PAGES	20
BATCH DELETING PAGES	21
DOWNLOAD PAGE PDF OPTIONS	23
REPLACING INDIVIDUAL PAGES	24
ADDING PAGES TO EXISTING ISSUES WITHOUT FILES	25
ISSUE-I EVEL NAVIGATION	26

## INTRODUCTION

The Newspaper feature of VITA Toolkit is a sophisticated way to display relationships between Publications and their contents.

With newspapers, we need to think in terms of structure and hierarchy—at least as we're building the records—so users can navigate and browse related content without losing a sense of which component part belongs with which related item.

The main differences in how we approach newspapers depend on your collection: are you working with full run newspapers? Or vertical files full of clippings?

With **full run papers**, you will be supplying issue after issue of the same Publication and uploading page image files for each of those issues. Once those are established, you can link index records to the individual pages too.

The relationships are a lot like family resemblances and hierarchies: Some information supplied about the Publication is inherited by the linked issues; some information about the Issue and Publication is inherited by the pages; some information from the Publication and Issue records is inherited by the Index records.

With collections of **clippings**, vertical file collections, or **BMD**s, ideally you would establish a Publication record for every Publication the clippings come from.

New records are created for each clipping: you can transcribe or scan the clippings themselves or simply capture as much index information as you can (title, issue date, proper names, Publication, page, column, etc.). The clippings will inherit the Publication title & location information. Each clipping record points the user to the proper page and issue date individually.

How the database thinks:

Publications have

Issues with

Pages with

Articles and BMDs

# NEWSPAPER PUBLICATION RECORDS

To get started with any Newspaper project, start by creating your Newspaper Publication "parent' record.

**Note:** All titles are assigned an ID number and are stored centrally so every agency can associate their index records to the entire list of titles.

## ADD PUBLICATION RECORD

Add a new Publication by entering a unique title on the Add/Edit Publication screen (below).

Remember to create a new Publication parent for every renamed Newspaper i.e. the *Weston Times*, the *Chronicle* and the *Weston Times-Chronicle* – tracking separate newspapers that were merged over time.



For full run newspapers the following fields are unique and **very important**:

## Secondary Media Type=Newspaper

In the Publication record, select **Newspaper** as the **Secondary Media Type** to activate the calendar display so that uploaded issue dates for are reflected in the calendar interface.

Click Update and you will see a new field "Role"

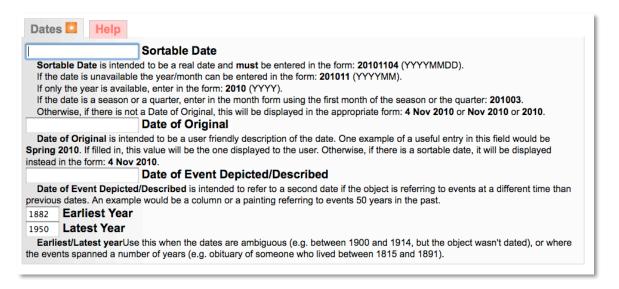
Role--Newspapers with Issues (only for News Plus & VINTA accounts)

Under **Role** select "Yes" if you will be assigning full **issues** to this Publication title. This activates the Add Issue uploader.



**Earliest & Latest Dates:** for a parent Publication, the date range is what is important—either the range of available issues in your collection or the range of dates that the Publication existed. In the case below, only the dates from 1882-1950 were digitized from microfilm and uploaded.

**Sortable Date:** use this field to inform the toolkit results "Sort by" option. Decide what consistent data you will input for your collection (earliest issue YYYYMMDD or another date).



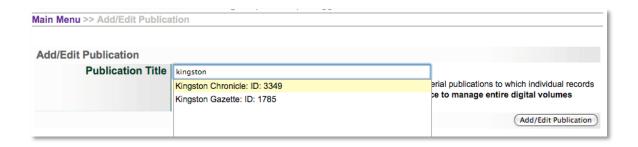
Complete the Publication record with other metadata and background information.

For more information on creating the Publication record with Descriptive, Geographic, Administrative metadata and assigning Groups or Links, see VITA Adding & Managing Records manual.

#### **EDIT PUBLICATION TITLE**

**Note:** you can only edit the Publication records that have been added by your agency. See "Title Management" (below).

On the main menu, click Add/Edit Publication and begin typing the Publication title. Select the one you want to edit from the drop-down list. On the following screens, you can edit the Publication record. See VITA manual for Publication record information (Data Management Manual, Section 5).



## PUBLICATION RECORD MANAGEMENT

Only the agency who adds a new newspaper title to the database can manage or alter that Publication record, e.g. associate a masthead to represent the Publication, add metadata that describes the Publication title. If you try to edit a Publication record that was added by another agency, you will receive the following message:

This record belongs to **Test 4.0.1 Database stuff**. You don't have rights to edit this record unless you "switch" to that agency.



#### ADD PUBLICATION MASTHEAD

Be sure to add a custom masthead to represent a specific Publication, create an image such as:







or use a full page scan of a masthead page:





Otherwise, the Publication record will be represented by a generic image tile like this:







Go to File/Tech data

Click "Associate files with this record"

Browse and select the chosen image from your hard drive

Choose Category: Thumb + Regular

Label—unnecessary

Change file size "Yes"

This image will be present in results sets and on linked article & BMD records

# **ADDING ISSUES**

Adding issue records with associated page image files and linking them to a parent Publication allows users to experience the newspaper in a natural presentation: Publication with linked issue dates that can be browsed or searched. Furthermore,

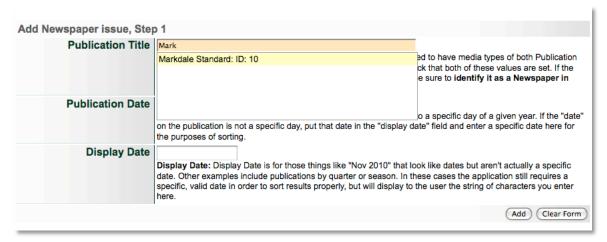
having Issue records means any index records can be linked to individual issues or pages.

## STEP 1: ADDING ISSUES

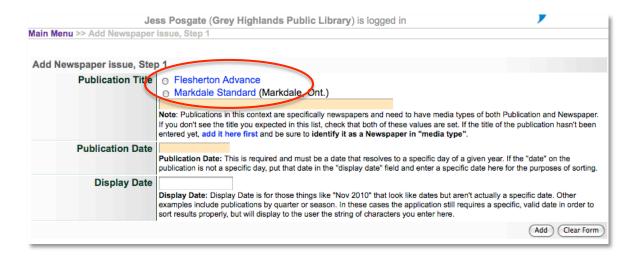
To add a new Issue, select "Add Issue" from the main menu

The information on the Add Issue screen is gathered to establish the relationship between an issue and it's parent Publication and fixes the issue date.

**Note:** You can only add issues to Publication titles managed by your agency.



If your agency has added Publication titles, they can be assigned as a quick list to display on this screen by any Level 2 or 3 user or the site administrator. (See Agency Management manual).



**Publication Title:** Start typing the Publication title and select the correct one from the drop down. The more of the title you type, the closer matches will be displayed. \*If the title does not exist, return to the main menu and Add Publication.

**Publication Date (of Issue):** Add the plain date of the Issue you are uploading. This will be translated into a machine-readable date, for sorting and ordering purposes, and into a human-readable display date, e.g. March 23, 1987

**Display Date:** Don't fill this in unless you are identifying an alternate display date like a quarterly, i.e. the issue is named "Spring 1984" not a specific date. Add a representative date for items like these in any case, i.e. for Winter 1984, use Jan 1, 1984; Spring 1984, Apr 1, 1984, etc.

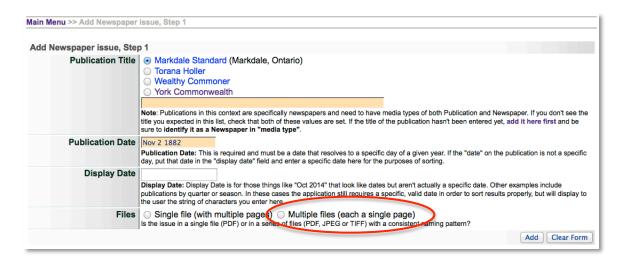
Why? Publication dates are available for multilingual transformation, whereas the Display date will remain as it is written.

## STEP 1: UPLOADING MULTIPLE FILES

For issues where every page is a separate file (JP2 or JPG + TXT or individual PDFs), choose the Multiple Files option. Then click "Add".

If you are uploading a single PDF, see the next section.

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



## STEP 1: UPLOADING A SINGLE PDF

Adding a single PDF for an issue means you do not have to strip the file names and, if it is a readable PDF, the text is automatically extracted for full text search.

On the Add Issue screen, choose the Publication title and add your Issue date.

Choose the "Single file (with multiple pages)" option and click "Add"

Add Newspaper issue, Step	o 1
Publication Title	Acton Free Press (Acton, ON) Acton Tanner (Acton, ON) Erin Advocate (Erin, ON) Georgetown Gemini (Georgetown, ON) Georgetown Herald (Georgetown, ON) Georgetown Independent (Georgetown, ON) Georgetown Leader (Georgetown, ON) Halton Hills This Week (Georgetown, ON) Halton Journal (Milton, ON) Independent & Free Press (Georgetown, ON) New Tanner (Acton, ON) North Halton Compass (Eden Mills, CN)  Note: Publications in this context are specifically newspapers and need to have media types of both of these values are set. If the title of the publication hasn't been entered yet, add it here first and
Publication Date	August 28 2012  Publication Date: This is required and must be a date that resolves to a specific day of a given y and enter a specific date here to the purposes of sorting.
Display Date	Display Date: Display Date is for those things like "Oct 2013" that look like dates but aren't actual application still requires a specific, valid date in order to sort results properly, but will display to the
Files	<ul> <li>Single file (with multiple pages)</li> <li>Multiple files (each a single page)</li> <li>the issue in a single file (PDF) or irra series of files (PDF, JPEG or TIFF) with a consistent name</li> </ul>

On the next screen, click "Add file" to select your PDF, then "Start Upload".

Click "Start upload" and when it reaches 100%, click the Continue.

## STEP 2: FILE NAMES

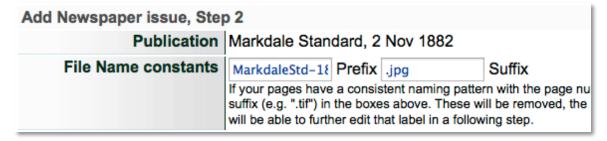
## File name

First, automatically assign page labels by "stripping" the common or "consistent" parts of the file names for all the issue files you are uploading. Do not include preceding zeros from the file name as these will be stripped out automatically.

E.g. the file names for my issue page files **all** include "MarkdaleStd-1882-1884-R2\_" (the prefix) and zeros that precede the page number; all the files are JPGs, so the extension or suffix is ".jpg".



On the Step 2 screen, include these constants in the Prefix and Suffix input boxes



## STEP 2: PAGE DISPLAY SIZE

The default page viewer will resize page files to 1000px wide for the smallest display. You can change this and determine whether to expand or reduce the display size for your newspaper pages. If you do not use the default, remember to be consistent from issue to issue.

If you have optimized the page images for your newspaper pages and want them to display at that size, display click "No" for the Change file size option; the default is "Yes" and allows the tool to automatically resize your image files to display at the selected size.

If you select "Yes", use the radio buttons to choose what display width will present best for the page content you are working with. Default is 1000px wide.

All Page files in VITA display with a pan-zoom viewer for JPGs and PDFs and a dynamic IIIIF viewer for JP2 files. For dense broadleaf or 6-8 column papers, it is better if the resize is larger than smaller so the zoom function has some depth for the end user or upload JPG 2000 files (recommended).

Change file size (page	● Yes ○ No
images only)	If "Yes" then what width?
	○ 800
	In general, make the width the smallest at which that page of the paper is legible. For with less than four columns
	800 pixels wide might be reasonable. Contemporary tabloid sized papers are quite legible at the 1000 width mark.  Older 6, 7 or 8 column papers might require a wider presentation. Remember that the larger the size the more the
	storage requirements and the bandwidth to deliver to your users.
	Add Clear Form

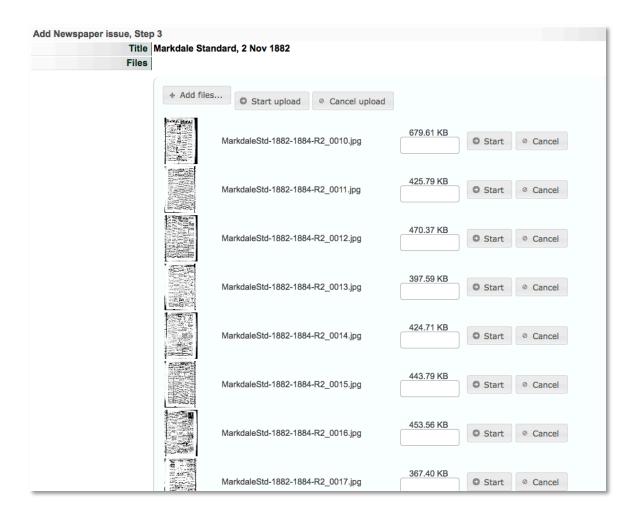
Then click "Add".

# STEP 3: MULTIPLE FILE UPLOADER

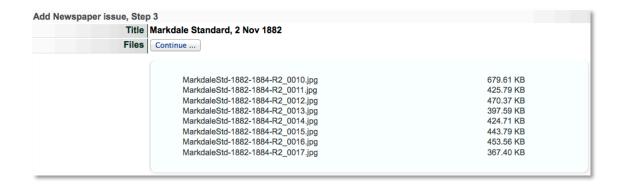
The next screen is for batch page file uploading.

Click "Add files" and select your page files.

Use Command or Ctrl + click to select multiple files from your hard drive.

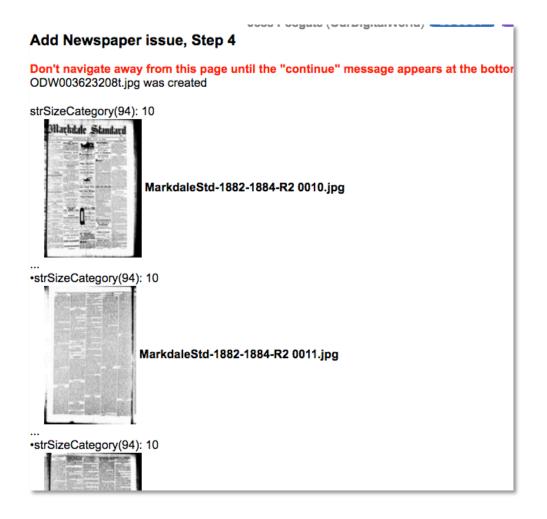


Click the "Start upload" button at the top of the screen to load all the files at once. Watch the status bars fill until all are 100% and the list collapses and a Complete button appears.

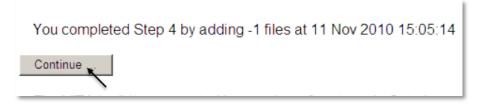


# STEP 4: MULTIPLE FILES PROCESSING & CONFIRMATION

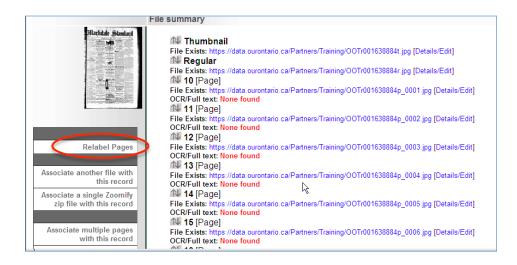
The following screen provides a real-time view of the page processing and snippets of any extracted text (readable PDFs only).



Once all the pages have been confirmed, scroll to the bottom of the screen and click "Continue..."

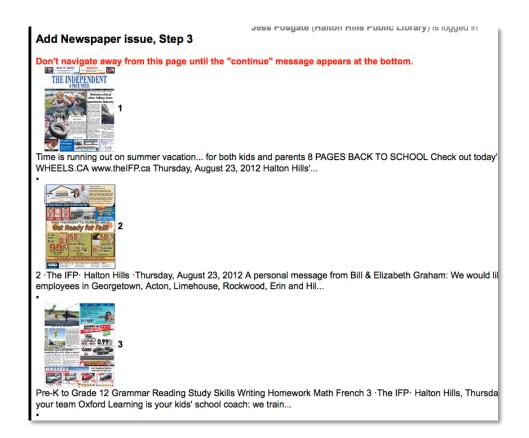


You will be returned to the File/tech data screen, where you can review the files you've associated. The labels are created from the unique numbers in your original file names; these can be customized in the Relabel Pages gallery.



STEP 4: SINGLE PDFS PROCESSING & CONFIRMATION

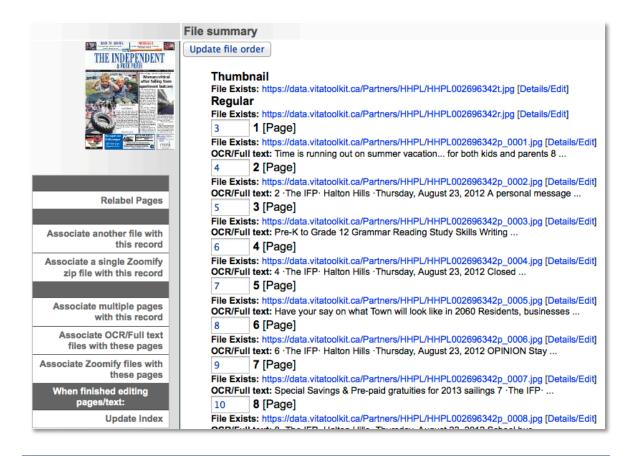
This screen shows a real-time view of the PDF splitting, page image page processing and text extraction (readable PDFs only).



Let the process run until a Continue button appears at the bottom of the screen. Click Continue.



You will be returned to the File/Tech data screen and see the individual pages, with snippets of the extracted text, and a single PDF file for optional download.

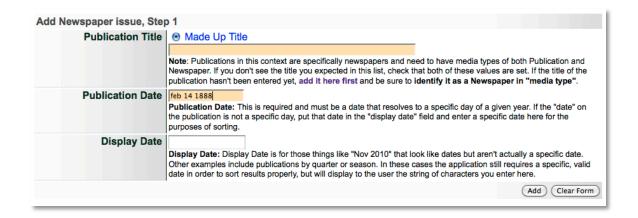


## MANAGING ISSUE FILES & CONTENT

## **EDIT ISSUE RECORD**

To adjust the record or files associated with an existing issue

Click Add/Edit Issue on the Main Menu Select the Publication title Enter date of the issue you want to edit Click "Add"

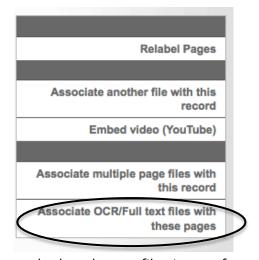


This will take you to the existing Issue record file/tech data screen, navigate through the record tabs to edit the record, add, delete or reorder page files, etc.

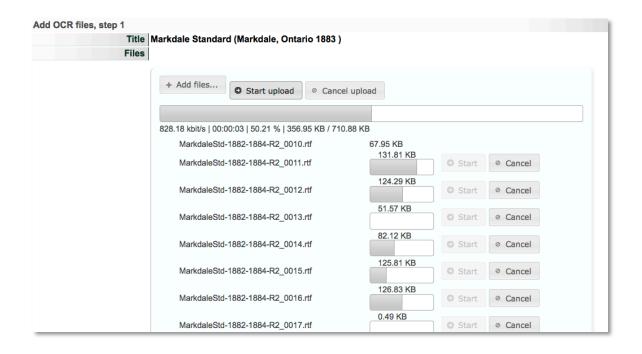
## ADDING FULL TEXT FOR MULTIPLE FILE PAGES

If you have OCR generated files for your newspaper pages, save the page content as .txt using **exactly** the same page file names as the image files e.g. MarkdaleStd-1882-1884-R2\_010.txt

Click on Associate OCR/Full text files with these pages



Browse and select the text files (.txt, .rtf or .pdf)



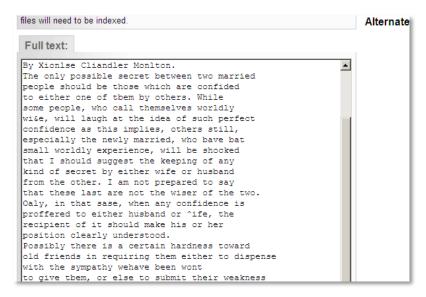
**Note:** Some PDFs and other proprietary text file formats (RTF) will "bulk up" your text content with wrapper code; whenever possible, save text content as plain text (.txt) or clean out the bulky content before uploading for better full text discovery.

E.g. This is the text extracted from a digitally-born PDF

```
Full text:
xœc'''c''Ra'dàøÙ
                                                                     •
$ € ™Á$ŒÏñ{% X= 30,1 38600äéMâOVVŽâq`)vbð4 •%Ý5 ½/þl 'l
€° `éÖq[fÆNő'ÊAçrr∏ó(ÇöOñ8'Úéüw&* * ^Pæ¾íb¯j2 =
E#ì§s'š"ŪÑjgŽ †ê7aÞ --€œ¬‡ç 5c ãýaþ_[Y,üí°tó öw¦µ¾«vÝs±Ü!
ù<q(à'÷Vy *Lz -Ô]0å·X$0{%?y H7Þ|w9.*¬ "ÎÈ"æ'xèŠ"^ œW. 0L\@
Äp= ĐæÚð0ïÞ°úÅÈ »5 7?‡(µ~0çS|/-- Ô°.`Hò3£-ñhŏUQÎÔWeBnJ> à|-
-¦C6 ) Æ 'ÄM2YB¿DÇ g-¿£TÇ
B'ÍÁ;¦ëÒüBô<~~|ĭË^Ú~...à5khTRýi¾ÿ| ÄÍüQØp<u°‡!Âçg=[ O¦Sq%
6jÕ'®4¬Î'lE;-SaŸx,á¼Õðsmž|ªå
¶~hë [üžÇi ëwª"÷
u7° = "O#ÊÍ'@XÌtú;V]/, žp,Ñ
Xû^ ãe ¾g-É
ÎOM"<ê_|±Œ
OúÏ 'Nî^¹¨ ¾pf /¬Öà• f ÚëC9 q ê/àpfWŠý ‡" ^SÛÜýìËSM…[]òÁt)
+ Ş<ÜÜİäm} من÷غ Âlãm $$ u...
                                    ®TÛ S®7Sœ#Lls"
Î9ðå¥l¿aï¶À|† >é³ô [¾Ÿ™õ2ú
Gîpè¿ ow\¶Œüè.ç? ¡ŠQ"PïË.m>s°ô^¥š4e'¥ÓD?4(ö% [ \J"(Ê)å |...
|n Mœ|9ŠvUÀG pzÝ ŸN[ ^+Đ'¹Jäê«,TðWí
                                             ï Á÷ ¡e/X•4Yøõ!
=^×LÏ &å©Ö ;f°X- ײ'~ÊĐX·"é<é.©R);R^ŪŒ`vx/ ôè,øÕ\ëŽpÊ m
Sxyü"Ô4;ñ°w2[?,°Ó' âÆZÁsl®lÞÈ;S‡©!Rà?$o\f:ŽD°¿kf#g
Èf©°K·2À[⊸l' 2Ó \PŽe•Ù8, ì D ØcöV[ê8»Q™âÑ#-¦
ïš4 K1¥Þ§ CVݾ ⊗Õ<£ií¼WV° 8RÒ%$f¦6À`v.Ã] S÷ÓÔ4Íž6ÑÎjH
lä=¶S" + 0 | 'ý6Jú:ñRäåëäñjä çq W+: jC...€...wúäñ x,)Lk42Íťæø,K ñ?
f2f֏L£`ì - ~-m"vš%®o¶' O".ß %M"J"';ú6'fœvåläl 4
±Ab'
```

Vs.

This text extracted from the same PDF by doing a simple "select all" from the PDF, then copying and pasting into a text editor like TextEdit or Notepad, saving as .txt and uploading as outlined above.



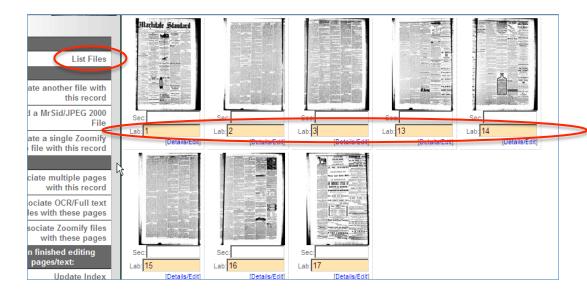
**Note:** You can edit full text on the details/edit screen for each file, but there is a character limit to that display that is often passed by newspaper page content. It is better to prepare your full text files and clean them before uploading page text files.

## BATCH RELABELING OF ISSUE PAGES

Click the Relabel Pages link

Add new label numbers for all pages.

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

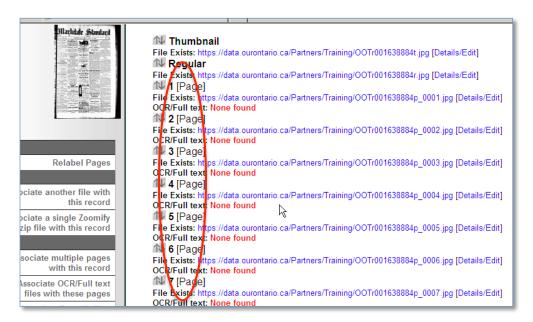


## Click **Update Labels** to commit changes

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.



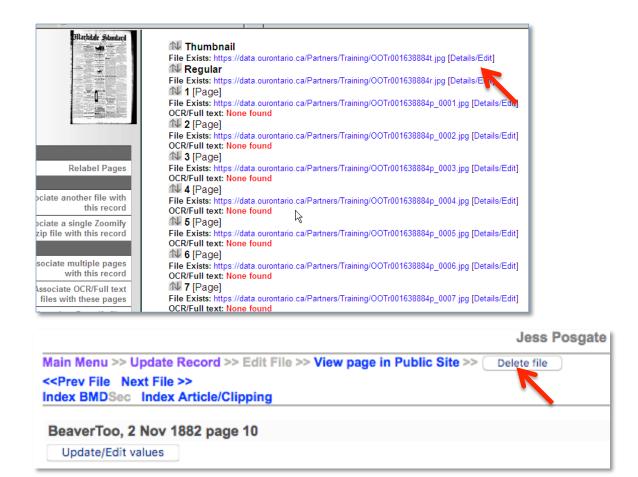
For issues with more than 20 pages, the page "order" will appear beside the page label in a numeral box. This order starts with the issue Thumbnail (assigned order 1), then the Regular (assigned order 2), and Page 1 (assigned order 3).

To reorder, either renumber the whole set and click update or, to insert one page between two others, use ##.5, i.e. if the masthead page is somehow buried in the rest of the pages, assign the masthead order number 2.5 and click update. The Masthead 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.

```
3
        1 [Page]
File Type: JP2
File Exists: https://data.vitatoolkit.ca/Partners/Test60/Test60000528401p_0001p.JP2 [Details/Edit]
OCR/Full text: P, "BoOoe | 1 L EC escTT Lo aws a the we * # t; f' T e '; § 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/Full 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 v 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 ...
        8 [Page]
File Type: JP2
File 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 issues 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).

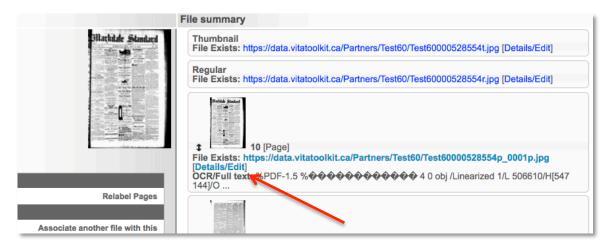
To batch delete pages, enter "D" in the numeral box for the files you would like to delete and click Update. This is an irreversible decision, so be certain you want those files removed from the record and the server.

## DOWNLOAD PAGE PDF OPTIONS

For page-level PDF download options, see the Agency Management manual.

## REPLACING INDIVIDUAL PAGES

In the event that you want to replace a page image file 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 file, and the download PDF file for that page (the first time, these PDFs are automatically derived from your original page file, and have no full text).

Replacing the page file will not disturb any lins 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.



## ADDING PAGES TO EXISTING ISSUES WITHOUT FILES

To add issue page images to a predicted issue date that has no files, go to the Main menu and choose Add/Edit issue



Select the Publication title from the quick list or start typing the title into the Publication title field and select the right title from the drop down menu that appears. Then add the issue date in the Publication date field in natural language. Choose whether you're adding multiple files or a single PDF that needs splitting.

Add Newspaper issue, Step	o 1
Publication Title	Acton Free Press (Acton, ON) Acton Tanner (Acton, ON) Erin Advocate (Erin, ON) Georgetown Gemini (Georgetown, ON) Georgetown Herald (Georgetown, ON) Georgetown Independent (Georgetown, ON) Georgetown Leader (Georgetown, ON) Halton Hills This Week (Georgetown, ON) Halton Journal (Milton, ON) Independent & Free Press (Georgetown, ON) New Tanner (Acton, ON) North Halton Compass (Eden Mills, ON)  Note: Publications in this context are specifically newspapers and need to have media types of be of these values are set. If the title of the publication hasn't been entered yet, add it here first and
Publication Date	August 28 2012  Publication Date: This is required and must be a date that resolves to a specific day of a given y and enter a specific date here for the purposes of sorting.
Display Date	Display Date: Display Date is for those things like "Oct 2013" that look like dates but aren't actua application still require
Files	Single file (with multiple pages)

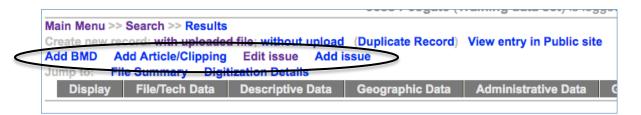
And proceed as for Add Issue (above)...

## ISSUE-LEVEL NAVIGATION

When you are in an Issue record, you'll see additional options in the navigation bar.

Use the links to

- a) Add a BMD record linked to the issue
- b) Add an Article or Clipping record linked to the issue
- c) Edit the issue you're currently in (will take you to the file/tech data screen)
- d) Add a new issue to the same publication (will take you to the add/edit issue screen with Publication pre-selected)



See the Indexing Articles & BMDs manual for more details about Index records.