Managing Your ResCarta® Collections

Using the ResCarta Toolkit version 6.0.2

June, 2016

The ResCarta® Toolkit is a product of the ResCarta Foundation, Inc., Onalaska, Wisconsin.

Using the ResCarta Toolkit, Version 6.0.2 June, 2016

This document provides a brief introduction to the tools included in the ResCarta Toolkit. More information is available from the ResCarta Foundation's website (http://www.rescarta.org); you can also contact the ResCarta Foundation via email: info@ResCarta.org

Table of Contents

Table of Contents....1
6.0.2 Additions / Improvements....3
Integrated OCR Engine....3
Background....1
Overview....2
Installing the ResCarta Toolkit....3
ResCarta Metadata Creation Tool....6
Subject Vocabulary Lists....12
Structural Metadata (Pagination, Sections, Chapters, etc.)....13
Using Metadata Creation Templates....16
Using the Metadata Creation Tool Carry Forward Feature....18
ResCarta Data Conversion Tool....22
Data Conversion tool “Object by file” mode:....24
ResCarta Textual Metadata Editor....26
ResCarta Audio Transcription Editor....29
ResCarta Collections Manager....33
For Oai_DC....40
For Dubin Core....40
ResCarta Indexer....41
ResCarta Checksum Verification Tool....43
Data Format Update Tool....44
Uninstalling the ResCarta Toolkit....46
Appendix:....49
The RcSystem.properties file....49
Sample JavaScript file....52
Adjusting tool memory usage....52
WARNING DATA LOSS!....53
Adding Addtional OCR Language Support....55

6.0.2 Additions / Improvements

- General
    - Improvements to the internal OCR engine by Tesseract
- Metadata Creation Tool
    - Support for multiple named notes. (page 11)
    - Support for subject vocabulary listings. (12)
- Data Conversion Tool
    - One click OCR conversion (page 22)
    - Creates **ALTO formatted output** for textural metadata. (page 50).
- Audio Transcription Editor
    - Export transcription to an external text file added to menu (page 32)
- ResCarta Indexer
    - Major rewrite of the Lucene Index
    - Support for searching all metadata elements at once.
    - Support for custom notes and alternate identifiers

.

Integrated OCR Engine

The ResCarta Toolkit, as well as having industry leading **Automatic Audio Transciption**, now provides direct Optical Character Recognition of images. The Toolkit now includes the Tesseract Open Source OCR Engine. Originally developed at Hewlett Packard Laboratories Bristol and at Hewlett Packard Co, Greely Colorado, all the code is licensed under the Apache License.

The Data Conversion Tool can process simple image files into full text searchable digital objects. The tool can also be used to reprocess existing image only ResCarta Formatted data to full text searchable digital objects.

Background

ResCarta® standards for scanning and naming conventions (brief overview):

Please review the following guidelines for digital documents to determine how to scan, photograph, and name your images; your data should meet or exceed these requirements for the tools of the ResCarta Toolkit to create a proper archive:

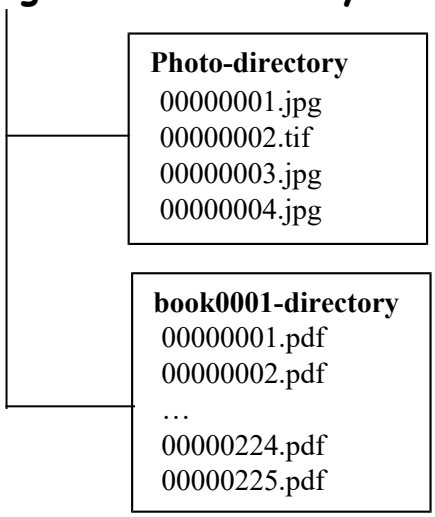
Scan book pages as any of the following:

- Bitonal, 600 DPI, G4 compressed TIFF files (suitable for pages containing mostly text, half-toned photographs, and etchings).
- Grayscale, 400 DPI, 8-bit, uncompressed TIFF files (suitable for pages containing mostly black-and-white continuous tone images).
- Color, 300 DPI, 24-bit, uncompressed TIFF files (suitable for pages containing spot color or color images).
- Save all images using a unique, sequential naming convention. For example, **00000001.tif**, **00000002.tif**, **00000003.tif**, etc.
- Save all images that represent a particular document in the same directory, and store all directories that represent your documents under the same "parent" directory. For example, the **.tif** images in the preceding example might be stored in a **book0001** (document) directory, which would then be stored in another directory (the "parent" directory in which all documents are stored).
- If book images are to be converted to text using an **external** OCR program, save output to PDF image over text (Searchable Image) as one file per each TIFF image file.

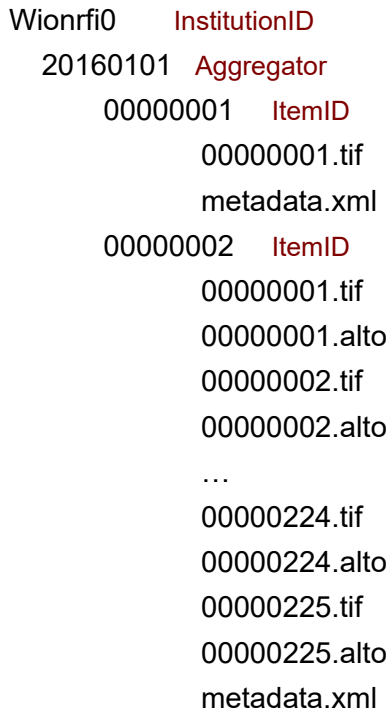
The ResCarta Toolkit also works with TIFF, JPEG, PDF (image only), PDF (image and text), PDF (Normal/Text) files and WAV audio files.

Note: The requirements here are a summary of the complete standards, which are available at the ResCarta Foundation's website: http://www.rescarta.org

Scanning Parent Directory



\RCDATA01



Overview

How the ResCarta Toolkit works with ResCartaWeb:

ResCartaWeb is a customizable web site with the ability to display and search an archive of ResCarta objects. The ResCarta Toolkit allows you to create, manage, manipulate, and check the ResCarta-formatted data accessed by ResCartaWeb, as well as add books or objects from your ResCarta archive.

Before you install the tools, make sure you have the correct Java Runtime Environment (see page 3). Also, uninstall any older versions of ResCarta tools you may have on your system.

There are eight tools in the complete ResCarta Toolkit. Using these tools will help you create a standard archive of data that you can view and search using ResCartaWeb or import into other web discover-display systems. You may need to use all eight or just a few of the tools, depending on what you want to accomplish.

For new images that you've digitized and are adding to your ResCarta archive, you will need to use tools that will convert images into ResCarta format. Specifically, you will need to use the ***ResCarta Metadata Creation Tool*** (see page 6) and the ***ResCarta Data Conversion Tool*** (see page 22).

For images that have not been OCRed, or for images to which you want to add searchable text (photographs, maps, etc.), you will need to use the ***ResCarta Textual Metadata Editor*** (see page 26) or the ***Data Conversion Tool*** so that textual metadata in these images will be searchable with ResCartaWeb.

For audio files that have been transcribed using the ***ResCarta Data Conversion Tool,*** or which you want to add searchable text, you will need to use the ***ResCarta Audio Transcription Editor*** (see page 29) to modify, edit or add text.

Whenever you make any change to your ResCarta archive (e.g., adding books or collections, images, changing the metadata, etc.), you must manage that change by using the ***ResCarta Collections Manager*** (see page 33) as well as index any changes by using the ***ResCarta Indexer.*** The ***ResCarta Collections Manager*** allows you to add and delete titles and edit metadata. The ***ResCarta Indexer*** (see page 41) insures that the new data you've added is added to the index so that when you use ResCartaWeb's search function, the new data will be searchable, along with the older data.

To verify the integrity of your data, you will want to use the ***ResCarta Checksum Verification Tool*** (see page 43). This tool, used after data has been converted to the ResCarta format, compares computed values of data to ensure accuracy. If you have created data with a previous version of ResCarta tools you may have to run the ***Data Format Update Tool*** to adjust the metadata.xml files for each digital object.

The tools included in the ResCarta Toolkit do not have to be installed on the same computer on which your ResCarta archive is stored, or on your ResCartaWeb server computer (regardless of whether the two are located on different machines). However, the computer on which you install the tools should have network access to the ResCarta archive, to transfer digital objects to the ResCarta archive after they've been converted to ResCarta format and edited as needed.

Note: The ResCarta tools and ResCartaWeb can be downloaded from the ResCarta Foundation's website: http://www.rescarta.org

Installing the ResCarta Toolkit

You can install any or all of the tools in the ResCarta Toolkit at any time. The tools are not required to be installed on the computer on which your ResCarta archive is stored, or on your ResCartaWeb server computer (regardless of whether the two are located on different machines), but the computer on which the tools are installed should have network access to the ResCarta archive computer. Before installing the ResCarta Toolkit, be sure that the computer on which you intend to install the toolkit meets the requirements for installation.

Note: While the ResCarta Toolkit may be installed on other operating systems, such as Apple® Mac OSX®, Red Hat® Linux® or Ubuntu; using the InstallRcTools_#_#_#.jar file, information in this document is specific to Windows systems. For information on installing the ResCarta Toolkit on other operating systems, please contact the ResCarta Foundation at info@ResCarta.org

Requirements for installing the ResCarta Toolkit:

- 1 GB RAM
- *Note*: 1 GB is the *required* amount of RAM for use with the ResCarta Toolkit; 4 GB or more RAM is recommended, especially for use with large audio data file or images (500,000+ images).
- Oracle Java™ Runtime Environment (JRE) 1.7.0 or greater
- *Note*: The latest version of the JRE is always recommended, and can be downloaded from the Java website at: *http://www.java.com/en/download/manual.jsp*
- *Tip*: To determine whether a machine already has a JRE installed, from the Windows **Start** menu, select **Settings**, and then select **Control Panel**. Control Panel opens. If Java is installed, it will appear in the list. Double-click **Java** to open the Java Control Panel, and on the **General** tab, click **About**. The About Java window appears. If the version of the installed JRE is *1.7.0 or greater*, you are ready to install the ResCarta Toolkit (see page 4); if not, you must first install the JRE (see next procedure).
- If you wish to convert Normal PDF files, you will need to install Ghostscript from Ghostscript.com or the ResCarta installer.

To install the Java Runtime Environment (JRE):

Note: It is strongly recommended that you exit all Windows programs before installing the JRE.

2. Open **Windows Explorer**, navigate to the JRE's executable installation file, and then double-click it.
3. The InstallShield Wizard appears, preparing to guide you through the Java platform installation process.

   Note: You can cancel the JRE installation at any time before it's complete by clicking **Cancel** in any of the installation windows. After you click **Cancel**, a Setup window opens to confirm that you want to cancel the Java installation process (click **Yes** to cancel the JRE installation, or click **No** to return to the installation).
4. The License Agreement window appears. Read the information in this window carefully, and then select the type of Java setup you want to use. You must accept the agreement to continue the Java installation.

   Note: The ResCarta Foundation recommends **Typical** setup for use with the ResCarta Toolkit.

   Click **Accept** to continue.

5. The Installing Java 2 Runtime Environment window appears for the specific version of the JRE you're installing, and Java installation begins. When installation is complete, the Installation Complete window appears.

   Click **Finish** to complete the installation process.

   *Not*e: You may be notified that you need to restart your computer; in that case, you will be given the opportunity to reboot immediately or at a later time. You may select either option, but be sure to reboot *before* beginning to install the ResCarta Toolkit.

To install the ResCarta Toolkit:

After you've confirmed that the JRE version 1.7.0 or greater is installed or have completed the installation process, you are ready to install the ResCarta Toolkit. You may install all the tools included in the ResCarta Toolkit or a selection. Complete the following installation sequence for the tools that you want to install.

Important! If you're installing an updated version of one or more of the tools in the ResCarta Toolkit, you should *uninstall* the older version of the tool(s) *before* installing the newer version. For more information on uninstalling these tools, refer to **Uninstalling the ResCarta Toolkit** (see page 46).

Note: It is strongly recommended that you exit all Windows programs before installing the ResCarta Toolkit.

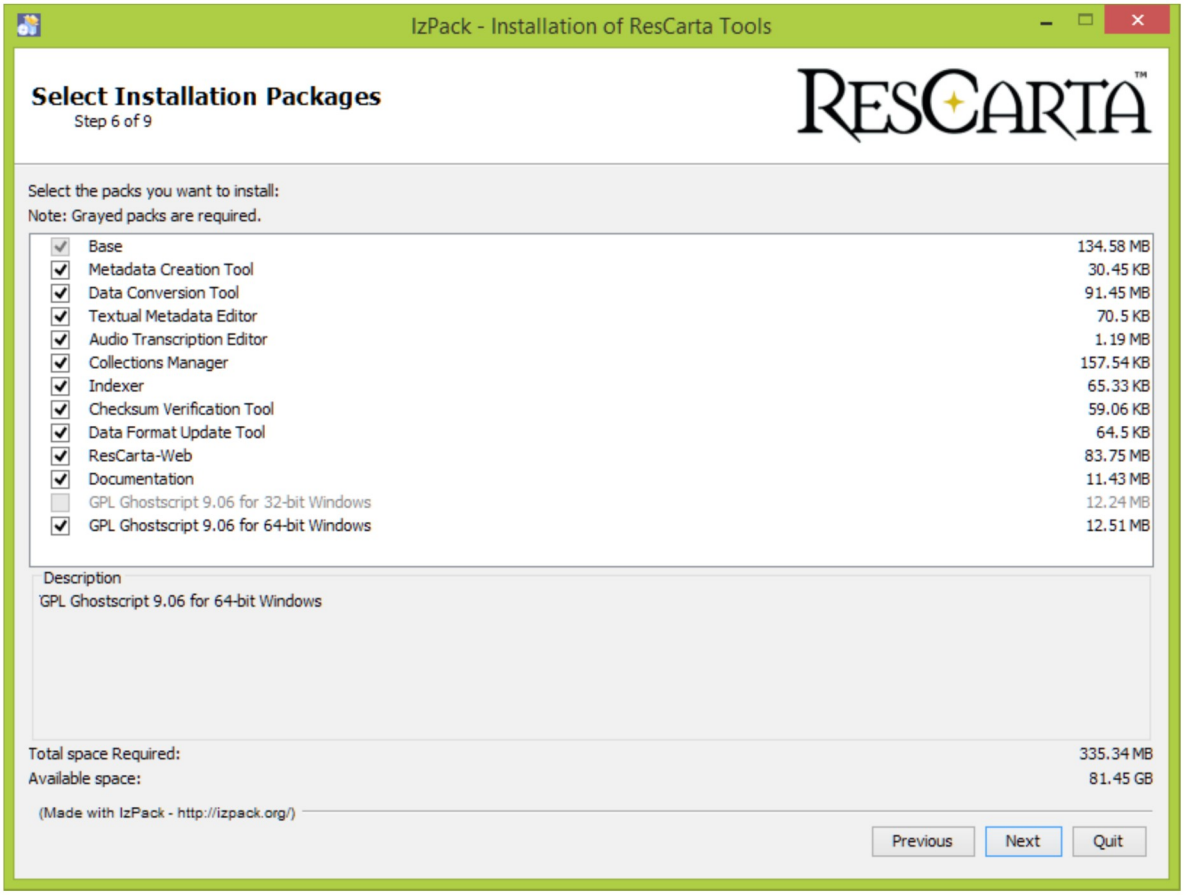
1. Open **Windows Explorer**, navigate to the Toolkit's executable installation file, and then double-click the file.
2. *Note*: The full file name of the executable file contains numbers and underscores immediately preceding the **.exe** extension. This portion of the file name represents the *version number* of the tool, and is not included in this reference document.
3. The IzPack Wizard appears, preparing to guide you through the installation process for the tool you selected.
4. *Note*: You can cancel the installation of any of the tools at any time before it's complete by clicking **Cancel** in any of the installation windows. After you click **Cancel**, a Setup window opens to confirm that you want to cancel the installation process (click **Yes** to cancel the installation of the tool, or click **No** to return to the installation).
5. The Setup Wizard's Welcome window opens. Read the information in this dialog and click **Next** to continue. You will be asked to accept the terms of the applicable licenses. Click "I accept" and next to continue.
6. The Setup Wizard's Select Destination Directory window opens. Select the folder in which you want to install the tools.

   Note: The default destination directory for the tools included in the ResCarta Toolkit is as follows:

   C:\Program Files\RcTools-#.#.# where #.#.# is the current version number.

   To select a different destination directory, enter the path to the folder in which you want to install the tools or browse to the desired installation location. On **Vista** and **Windows 7,** you may install into your personal directory since the installer is not allowed to install into the \Program Files area, unless you are installing **AS** the **ADMINISTRATOR**. Click **Next** to continue.

6. Select installation packages from the next dialog box. Select the tools you wish to install on your computer system by clicking on the checkboxes. Press the Next button to continue.



7. The Setup Wizard's Summary window opens, click next and installation begins. The installation progress bars will show the individual and overall progress of the installation. After completion of the installation, the installer will ask if you want start menus and if the software should be set up for your user or all users of this system.

   ***Note*:** The default Start menu folder for all of the tools included in the ResCarta Toolkit is **ResCarta Tools-#.#.#**

   To select a different Start Menu folder, enter the name of the folder in which you want to display the tool's shortcuts, or browse to the desired display location.

   If you choose to install **Ghostscript** you will be presented with a separate installation program. You must agree to a separate license, and select a destination folder. Uncheck the "Show Readme" box and click the **Finish** button to complete this installation.

   Click **Next** to continue.

8. When installation is complete, the Setup Wizard's Completing window opens.

   Click **Done** to close the Setup Wizard window and exit setup. It is important that the last button be the **Done** button to assure that all steps of the installation have been completed.

After you've completed the installation of the eight tools, the ResCarta Toolkit is installed and ready for use.

ResCarta Metadata Creation Tool

Use the **ResCarta Metadata Creation Tool** to create *object metadata* (information about your image files that's stored with them in a digital format). Any new, digitized object that you create for inclusion in your ResCarta archive needs object metadata in order to become part of a standard ResCarta archive.

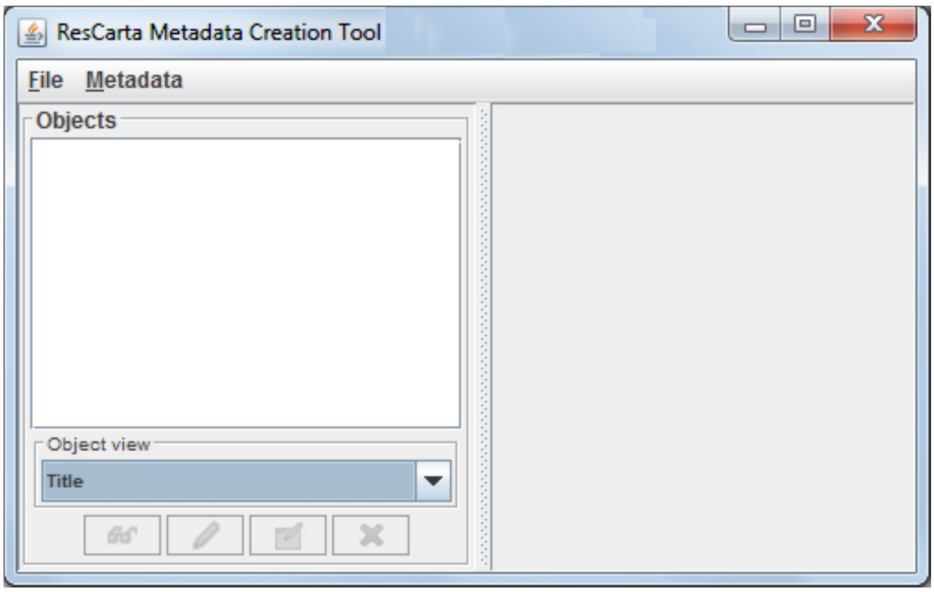
Let's assume we have a directory containing image files made by scanning the pages of a book. Here we would want the images of the pages to be considered as a single "Document"

To create metadata in "Object per Directory" mode:

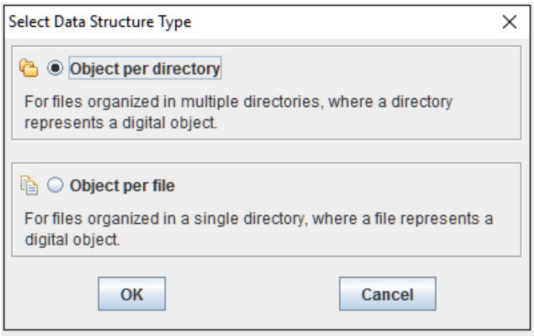
1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4, and then select **ResCarta 1. - Metadata Creation Tool**.

Note: The version number of the **ResCarta Metadata Creation Tool** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

The **ResCarta Metadata Creation Tool** appears.

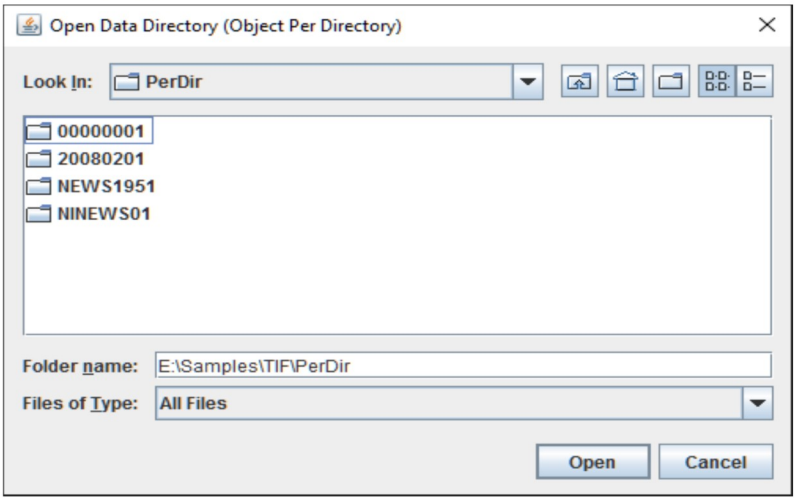


2. From the **File** menu, select **Open image data directory**.

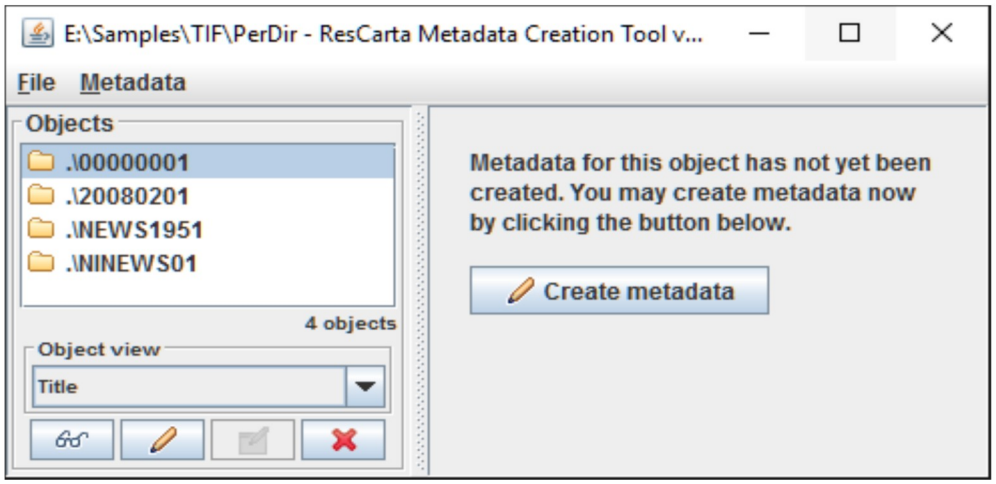


3. From the **Mode** dialog, select **Object per directory,** press the radio button and click OK.

The Select directory dialog appears.



4. Enter the full path or browse to the directory in which the images for which you want to create metadata are stored and click **Open**. In this example we have selected a directory called PerDir, which contains four directories, each with the image files of a particular newspaper.
5. The Image data directory opened window appears to inform you how many subdirectories (representing documents) were found within the parent directory you selected. Click **OK** to continue.
6. The **ResCarta Metadata Creation Tool** now displays a list of documents within the parent directory you selected (in the left pane) and displays metadata associated with the current document (in the right pane).

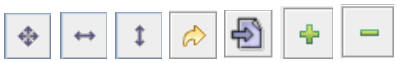


Note: Documents are listed by their directory location and file name until **Title** metadata is assigned. **Documents** for which no metadata has been created are indicated with an folder icon: **Documents** for which metadata exists are indicated with a book icon:

Tip: You can view the entire directory location and file name (for documents for which no metadata has been created) *or* title (for documents for which metadata already exists) of any item in the **Documents** list by briefly pausing your mouse pointer over the item; the complete directory location and file name or title appears next to your mouse pointer.

7. To select a document, click its name in the **Objects** pane list. Any available metadata for that document is listed on the right.

9. To view images for the current document, click **View images**.  An image viewer appears, displaying the first image in the current document.

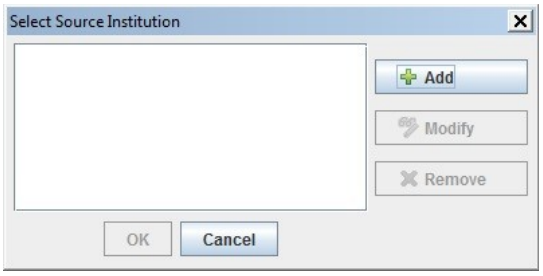


These options are, respectively, **Best fit** (displays the image at the height and width that are proportionally sized to fit the viewer), **Fit to width** (displays the image at the largest width that fits viewer), **Fit to height** (displays the image at the largest height that fits the viewer), **Rotate** (rotates the image 90 degrees to the right), **Go to page** (opens a dialog for you to select a specific image in the document to view), **Zoom In** (enlarge the view) and **Zoom Out** (reduce the view).

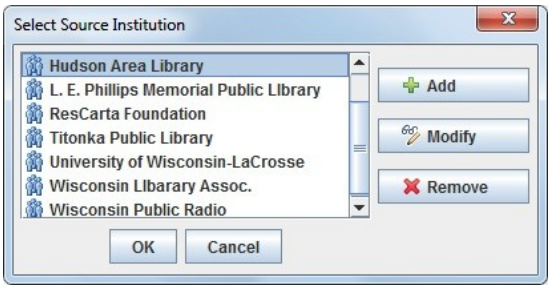
When you're finished viewing the images within the document, click the viewer's *close window* button. 

10. To create new metadata for the current document, click **Create metadata**. 

or click the **Pencil icon** on the bottom of the Left Pane.  This will open the Institutional ID dialog.



To add a new source institution, click **Add**. The Add institution dialog appears. Enter the **Id** for your source institution (8 character alpha-numeric), enter the full **Name** of your source institution (optional), and click **Finish**.

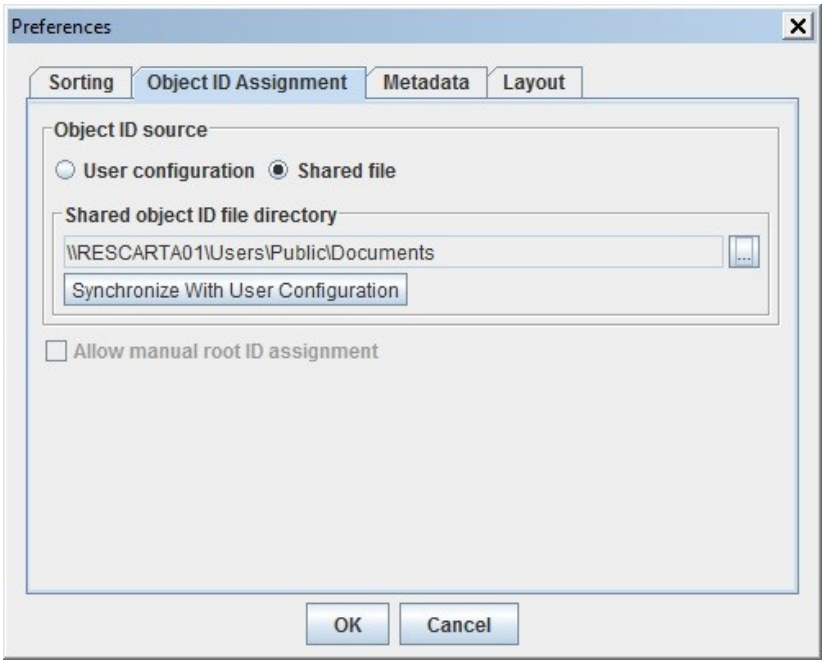


Important! The source institution identifier *must* be eight characters (letters and/or numbers) in length. You can use multiple source institution identifiers with your ResCarta collections, but a source institution identifier is *required* for each document. To avoid collisions when sharing data with other institutions, this identifier must be unique to your institution. If your institution has a code in the **MARC Code List for Organizations**, we recommend that you use that code, padded to the right with zeroes for a total of eight characters, as the source institution identifier. For example, the MARC code for the New York Public library is **NN**; therefore, the recommended source institution identifier for the New York Public Library is **NN000000,** ResCarta's MARC code is **WiOnRFI**, which is normalized to wionrfi0. See the Marc Institutional codes at http://www.loc.gov/marc/organizations/

To modify an existing source institution, select the institution you want to change and click **Modify**. The Modify institution dialog appears. Change the **Id** and/or **Name** of the source institution as needed and click **Finish.**

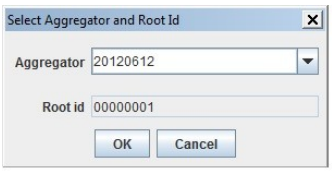
Note: If metadata already exists for the current document, the **Create metadata** button is not available; use the **Modify metadata** button to modify the existing metadata.

Shared or local Identifiers: From the **File** menu, select **Preferences** and click on the Object ID Assignment tab. This will set the location of the file to store your identifiers. It can be a local user or a network shared location. If you have started using ResCarta tools alone and want to share the information with a team, select **Shared File**, enter or browse for a network location where you and your team have write permission and then press **Synchronize with User Configuration** this will allow you to share your existing Institution Identifiers and coordinate your team.



To remove a source institution, select the institution you want to remove and click **Remove**. The Confirm institution removal dialog appears. Click **OK** to remove the source institution; click **Cancel** to cancel the institution removal process.

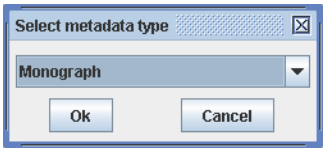
The Select aggregator and root id dialog appears. Enter the **Aggregator** you want to use and click **Ok**. Aggregators create subdirectories to limit load on operating systems. Below is an example of using the date as a non-repeating aggregator.



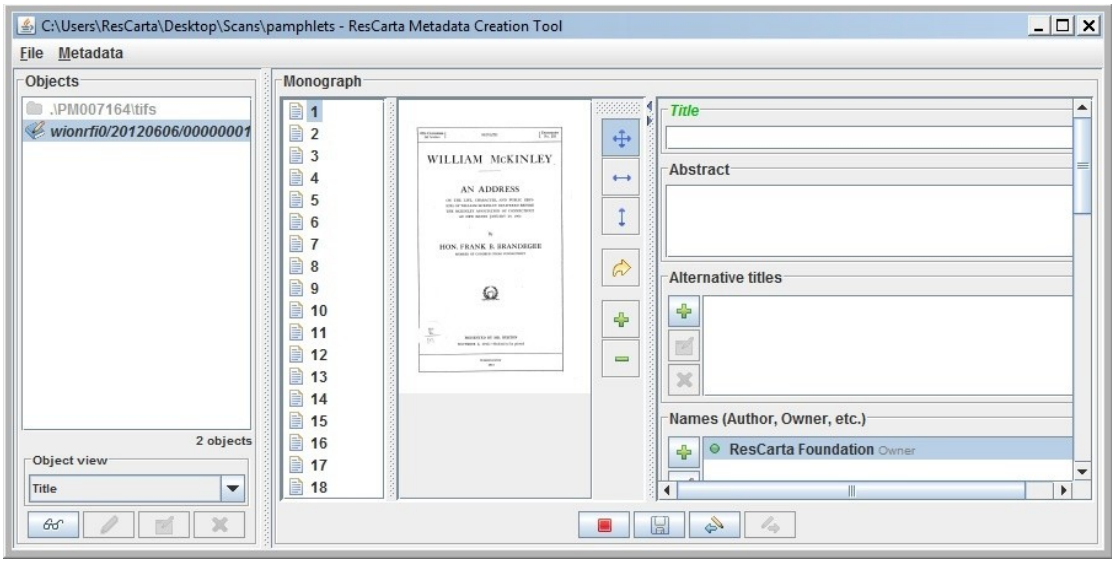
Important! The aggregator and root identifier *must* each be eight characters (letters and/or numbers) in length. An aggregator and root identifier are *required* for each document in your collection. The tool default is a number padded to eight characters. You can use the current date as a simple aggregator (e.g. 20120612)

Note: Collectively, the source institution identifier (**Id**), aggregator (**Aggregator**), and root identifier (**Root id**) that you just specified comprise the *object identifier*, which must—in combination—be unique for each object, so that it can specify the location of that object. This directory structure (***institution identifier/aggregator/root identifier***) is referenced throughout the ResCarta Toolkit.

- The Select metadata type dialog appears. Select the *type* of the current document (**Monograph**, **Serial monograph**, **Serial, Photo, Newspaper or Audio**) and click **Ok**.



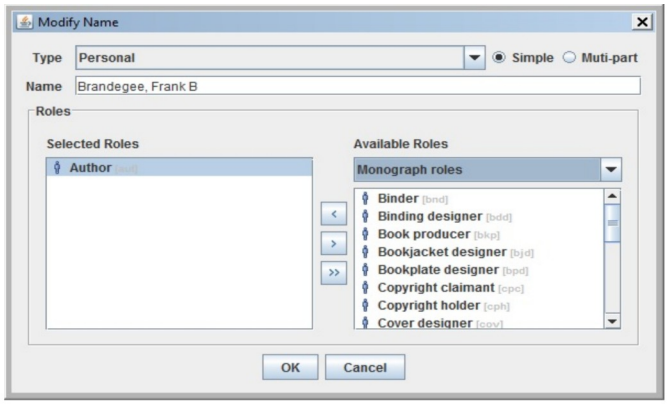
- The Create monograph metadata, Create serial monograph metadata, Create serial metadata, Create Photograph metadata, or Create Newspaper metadata dialog appears depending on the type of metadata you selected for the current document; the Create monograph metadata dialog is shown here:



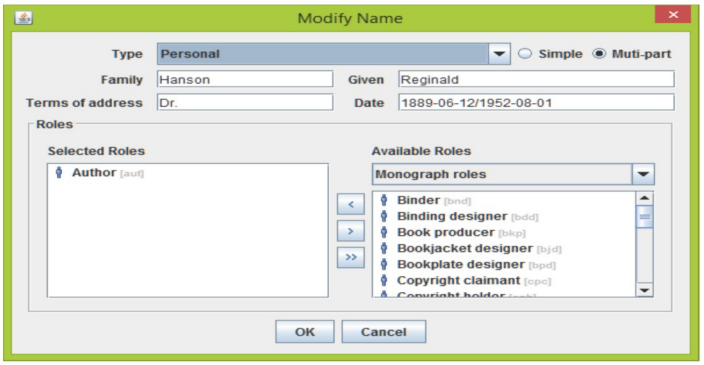
Note: A **Title** is *required* to add metadata to a document; if you are adding serial monograph or serial documents, a **Volume** is also *required*. The ResCarta Foundation recommends completing as much metadata as possible for all documents in your collection.

Information that you can enter includes: **Title**, **Volume** (for serial monograph or serial documents only), **Abstract**, **Alternative titles**, **Names(Author**, **Owner**, etc.), **Type of resource**, **Genre**, **Publisher name**, **Place of publication**, **Publication date**, **Capture date**, **Language**, **Size**, **Note**, **Subjects**, **Acess Restrictions, Alternate Identifiers**, **Issues**, **Sections** and **Pages**.

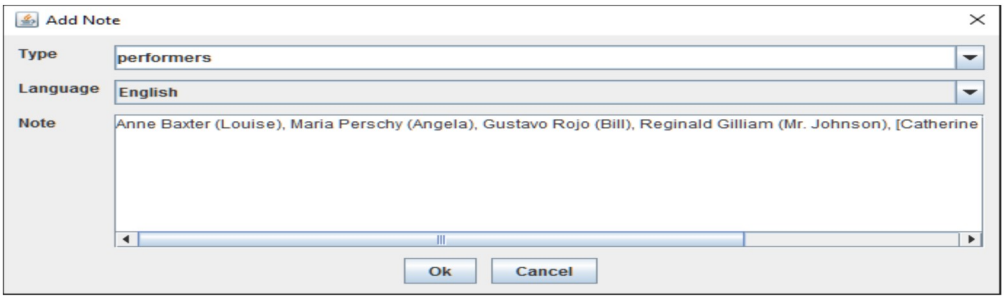
Note: Creating, changing, or removing metadata for **Alternative titles**, **Names**, **Languages, Notes,** and **Subjects** requires the use of the **Add**, **Modify**, and **Remove** buttons, respectively. These buttons appear directly to the left of the section for which they're used. Adding **Names** calls the **Add Name** dialog show below.



The Names dialog allows for multipart and simple formats and can restrict a listing to the material type (e. g. Monograph roles) or a complete listing of roles available.

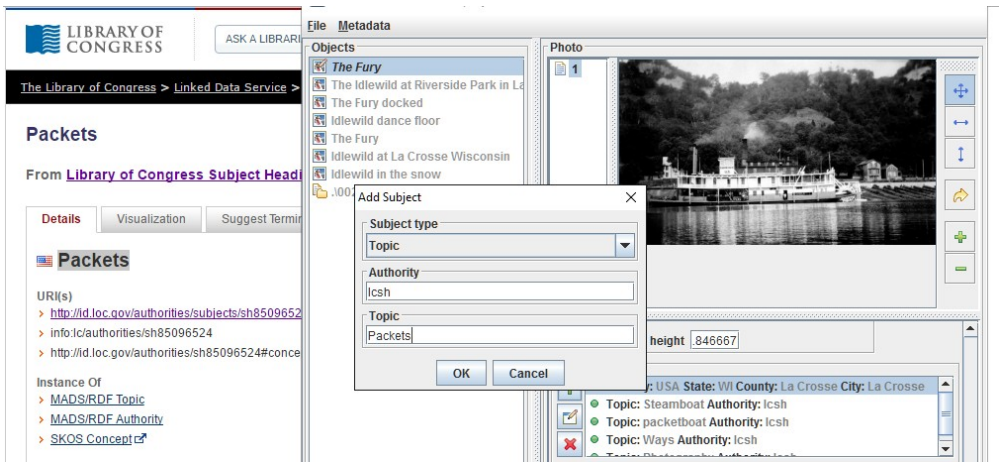


The Notes field allows for multiple fielded notes with attributes.



Subject Vocabulary Lists

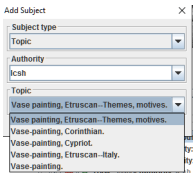
Topic lists can be added to the dialog for creating consistent subject headings for your collections. You can type single topics into a dialog. Click the “PLUS” icon in the Subject box to open the Add Subject dialog. The default will open with “Topic” selected. You can type in an Authority and a Topic into this dialog box. The example below shows the Authority as “lcsh” and the Topic as “Packets”.



If you want to have a fixed listing of selected terms from a particular authority, you can add a list to your .RcTools/vocabulary/topic directory. The . RcTools/vocabulary/topic directory is created when you first use the Metadata Creation Tool. It contains a sample “WRLS” authority file. The sample local authority file was created by taking a listing of terms from the following url.

http://www.wrlsweb.org/wp-content/uploads/2015/10/Subject-vocabularyWRLSECHOProject.pdf

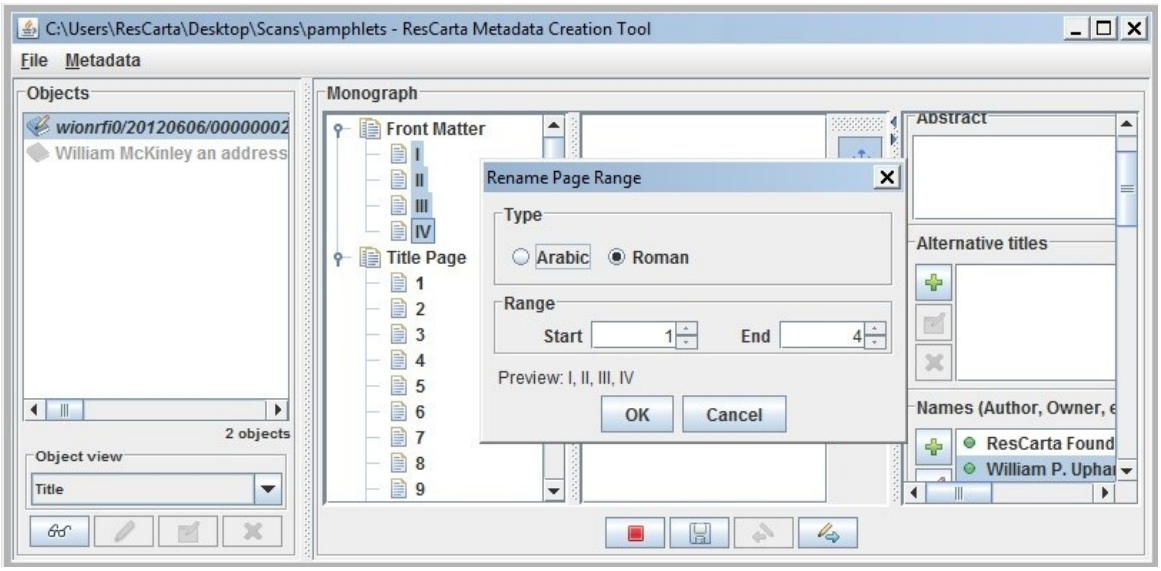
The terms were copied from the PDF file and placed into a text file called “WRLS.txt”. The name of the authority “WRLS” is derived from the WRLS.txt file in the .RcTools/vocabulary/topic directory. To create your own vocabulary listing, first create a named file for the authority like “lcsh.txt” in the . RcTools/vocabulary/topic directory. Then add a list of terms you wish to use for your vocabulary with one term per line. The resulting pull-down will be created.



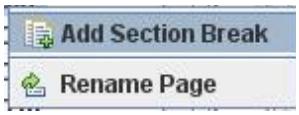
To remove the sample “WRLS” list from your system, delete the “WRLS.txt” file from your local .RcTools/vocabulary/topic directory. This listing is on a PER-USER basis. Each user will have their own private vocabulary directory. If you want to have a shared standard file across all users you should publish the “AUTHORITYNAME.txt” file for users to copy into their respective .RcTools/vocabulary/topic directory.

Structural Metadata (Pagination, Sections, Chapters, etc.)

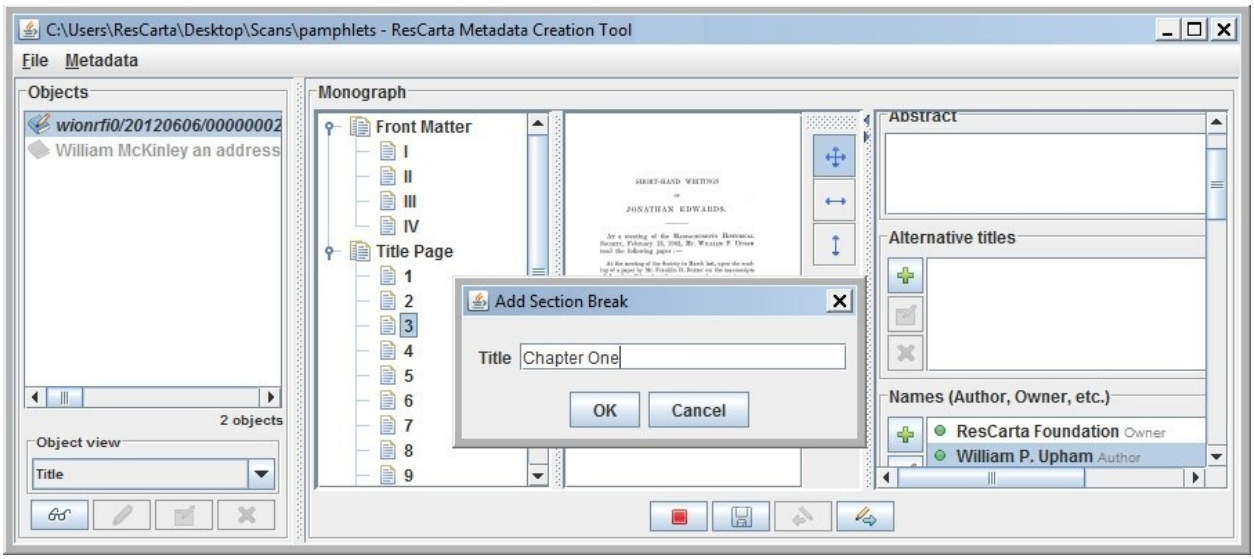
An image viewer is included in the **Pages** section for convenient page reference to **Modify page metadata right click a page name** this accesses the Modify page metadata dialog for renaming pages. You can select a range of pages and rename them to an Arabic or Roman range.



Tip: To create section breaks like chapters, issues, and divisions or to denote an index; right click a page name and select "Add Section Break" from the dialog.



You can then add a tag name to the page for a section break. This text is stored in the logical structure of the object. This is the same method used for creation of issue breaks when creating metadata for Serial objects.



.

Tip: The **Publication date** and **Capture date** must be specified in one of the following, three ISO formats, where **YYYY** is the year (four-digit format), **MM** is the numeric value of the month (in two-digit format), and **DD** is the day (also in two-digit format): **YYYY-MM-DD**, **YYYY-MM**, or **YYYY**. A date range is formatted : **YYYY-MM-DD/YYYY-MM-DD**, or **YYYY/YYYY**

- Click  to save your changes; or click  to close the metadata dialog **without** saving your changes.

11. To modify existing metadata for the current document, click **Modify metadata**  button.

    Note: If no metadata exists for the current document, the **Modify metadata** button is not available; use the **Create metadata** button to create new metadata.

12. To remove all metadata from the current document, click **Remove metadata**.  The Confirm metadata removal window appears. Click **OK** to remove all metadata from the current document; click **Cancel** to cancel the metadata removal process.

    Note: If no metadata exists for the current document, the **Remove metadata** button is not available.

13. When you are finished creating metadata, click the **ResCarta Metadata Creation Tool**'s *close window* button.  The confirm exit window appears. Click **OK** to exit; click **Cancel** to return to the **ResCarta Metadata Creation Tool**.

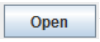
    Note: Document metadata is saved in the **metadata.xml** file, in the directory in which the image files for that document are stored. This metadata is saved in the METS (Metadata Encoding & Transmission Standard) format. For more information about METS, refer to the METS website at the Library of Congress: http://www.loc.gov/mets/

To create metadata in "Object per file" mode:

From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **1. - Metadata Creation Tool**.

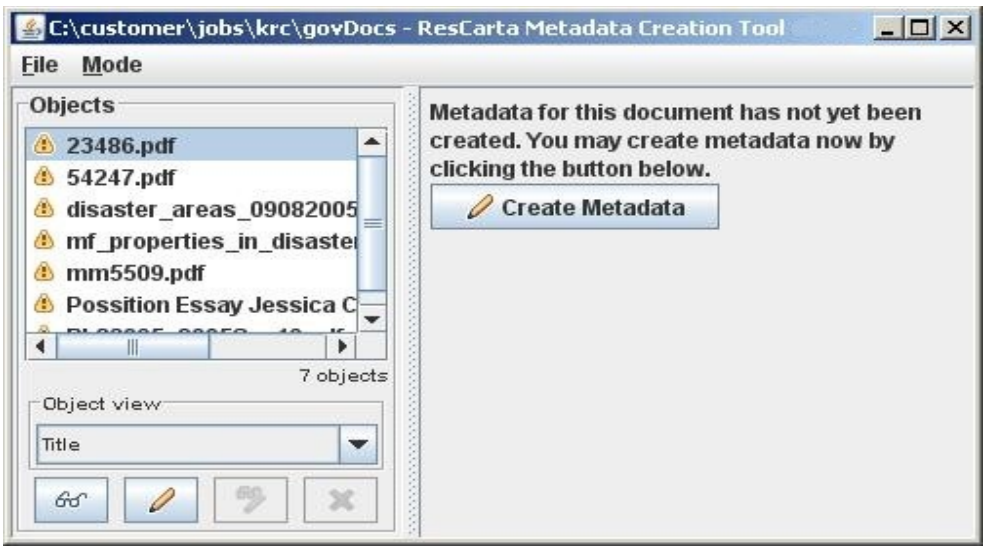
The **ResCarta Metadata Creation Tool** appears.

2. From the **File** menu, select **Open image data directory**.

3. From the **Mode** dialog, select **Object per File.**
4. The Select directory dialog appears.
5. Enter or browse to the directory in which the images for which you want to create metadata are stored and click **Open** 

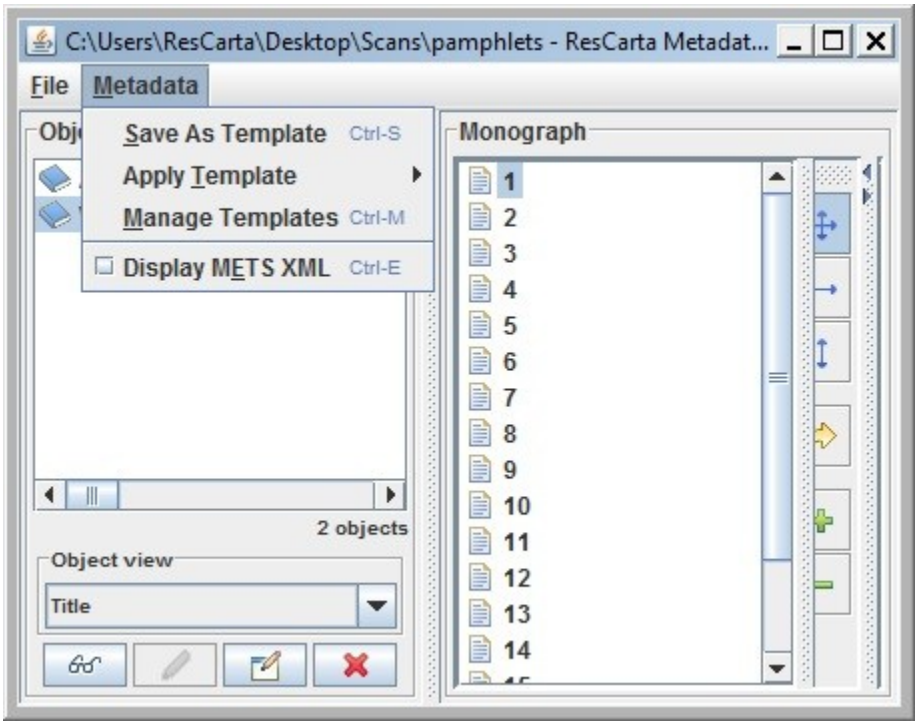
button.

This will allow you to add a metadata.xml file for each file in the directory. You can apply individual metadata elements to each file. This is primarily used for photographs, audio or multipage PDF file conversion. Proceed as with step 5 above to produce metadata files.



Using Metadata Creation Templates

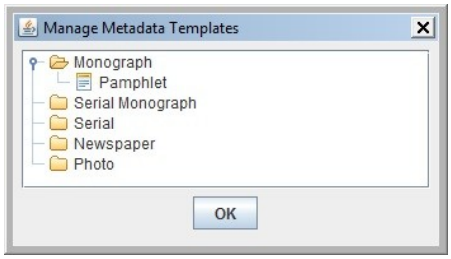
If you are creating metadata for more then one item with similar metadata, you can use the **Metadata** menu to access the templates menu items.



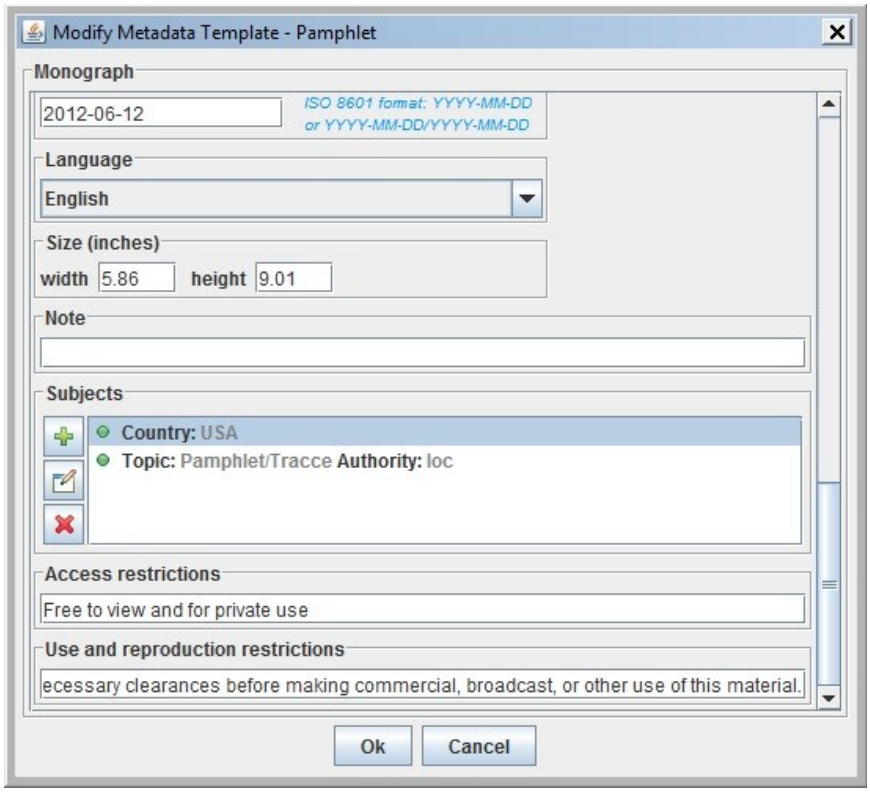
After selecting and opening a directory structure, you will be able to enter metadata into the metadata creator form. After any field has been entered, you may save the metadata to a template file to be reused. From the **Metadata menu** select **Save As Template** and give the metadata a name. The Metadata Creation Tool will save the template in the appropriate format area.

To use a saved template select an object from the objects pane, then open the **Metadata menu** and navigate to **Apply Template**, then to the format (Monograph, Serial, Photograph, etc.) and select a saved template to apply.

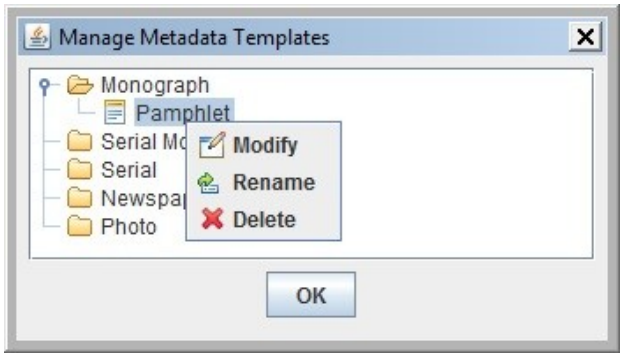
You may edit the contents of saved templates by opening the **Metadata menu** and selecting Manage Templates. This will open the manage templates dialog as shown below.



Double click on an existing template icon and the template form with existing template data will open. You may add, modify or delete any field on the form as you would normally during metadata creation.



To rename or delete a template, right click the template icon with your mouse and select the **Rename** or **Delete** functions.

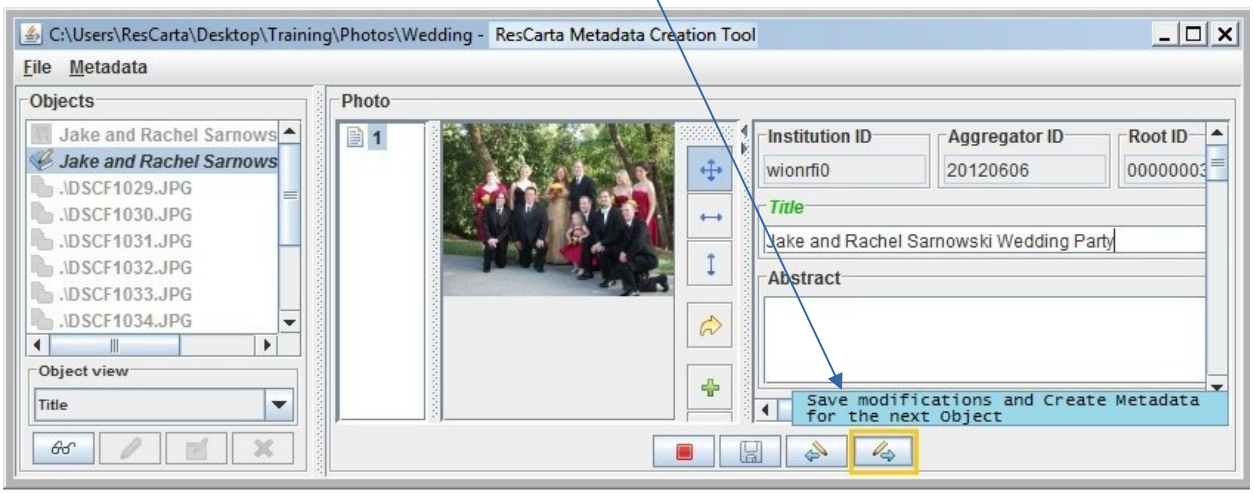


Using the Metadata menu, you may also turn on a display of the METS XML data that the Metadata Creation tool is writing. This will open a tabbed display in the metadata form pane. You may switch between the tabs as you edit the metadata creation form to see the changes being made to the XML file. You may close the XML display by clicking the x on the right corner of the METS XML tab.

Using the Metadata Creation Tool Carry Forward Feature

There are times during the entry of metadata where almost all of the metadata elements are nearly the same. For instance when one is entering metadata for photographs of homes or businesses on a particular street, most of the metadata will be similar if not the same. In this case it would be helpful to use the Carry Forward Feature.

Once you have entered information into the metadata form, the "Save modifications and Create Metadata for the next Object" button (see image below) on the bottom of the metadata form pane will become available. When you click this icon all metadata fields having data will be carried into the form of the next object in the object pane. In the case of the street photography, you may only have to adjust the title containing the address. Or for event photography you may only have to replace the title.



Once you have used added metadata beyond the first item in an object list the **"Save modifications and Create Metadata for the previous Object"** will become available.

As you click the **"Save modifications and Create Metadata for the next Object"** the metadata will be automatically saved. You may stop at any time by clicking the STOP button. Or click the save button to save your data and stop the editing mode.

The **Carry Forward Feature** can be combined with the use of stored metadata **Templates** to greatly reduce repetitive entry in the creation of metadata for your images.

ResCarta Metadata Creation Tool – Preferences

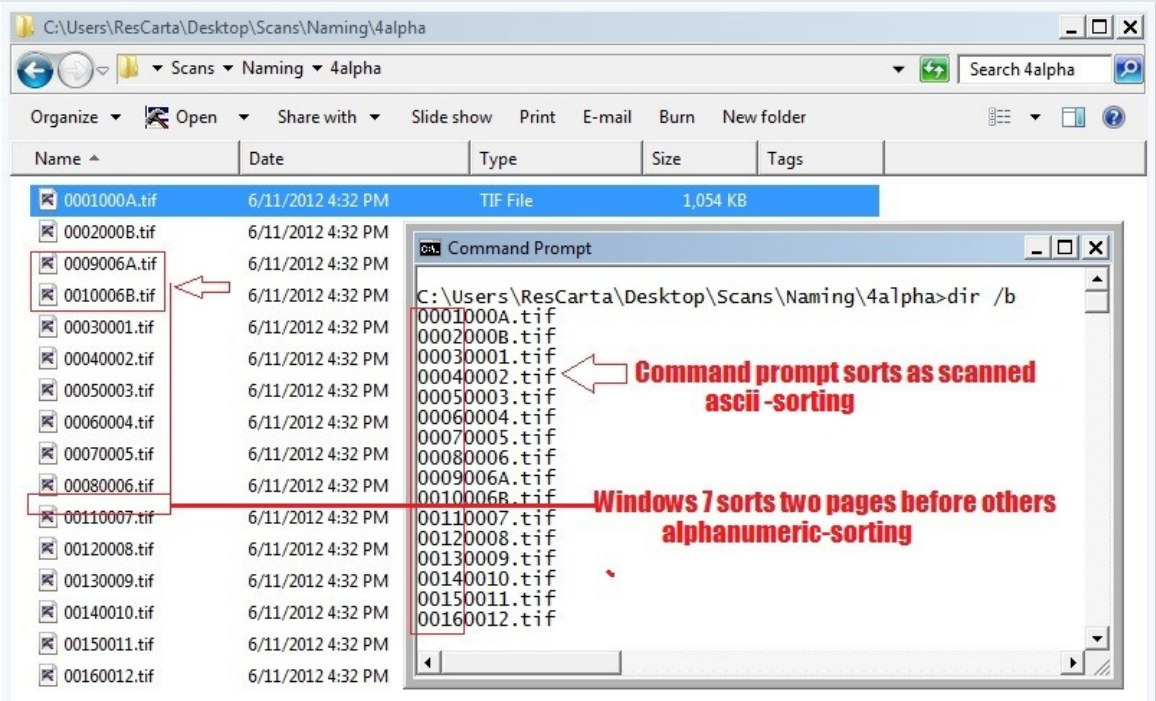
There are multiple custom settings for the Metadata Creation Tool which can be accessed from the File/Preferences menu. The preferences dialog has four tabs.

Sorting Tab

This dialog will set the default sort used for the image file names in your working directory. The tool uses ALPHANUMERIC sorting by default, which is used by most modern operating systems. You can change this setting to ALPHABETIC sorting or UNICODE sorting. The dialog contains examples of how the sort would handle mixed alpha and numeric names.

The suggested naming convention for scanning is to use eight character padded sequential names (e.g. 00000001, 00000002, 00000003, 00000004 etc.). Some scanning programs or vendors do not allow this simple format. Previously ResCarta used the "Making of America" (MOA1) file naming convention which used four characters for the sequence of the image and four characters for the page name. To work with files scanned to the MOA1 conventions; change the default sort to ALPHABETIC to order the pages properly.

Different sorting methods can produce page listings in unwanted order, as can be seen in the screen shot below from a Windows 7 system and the same system using the cmd prompt.

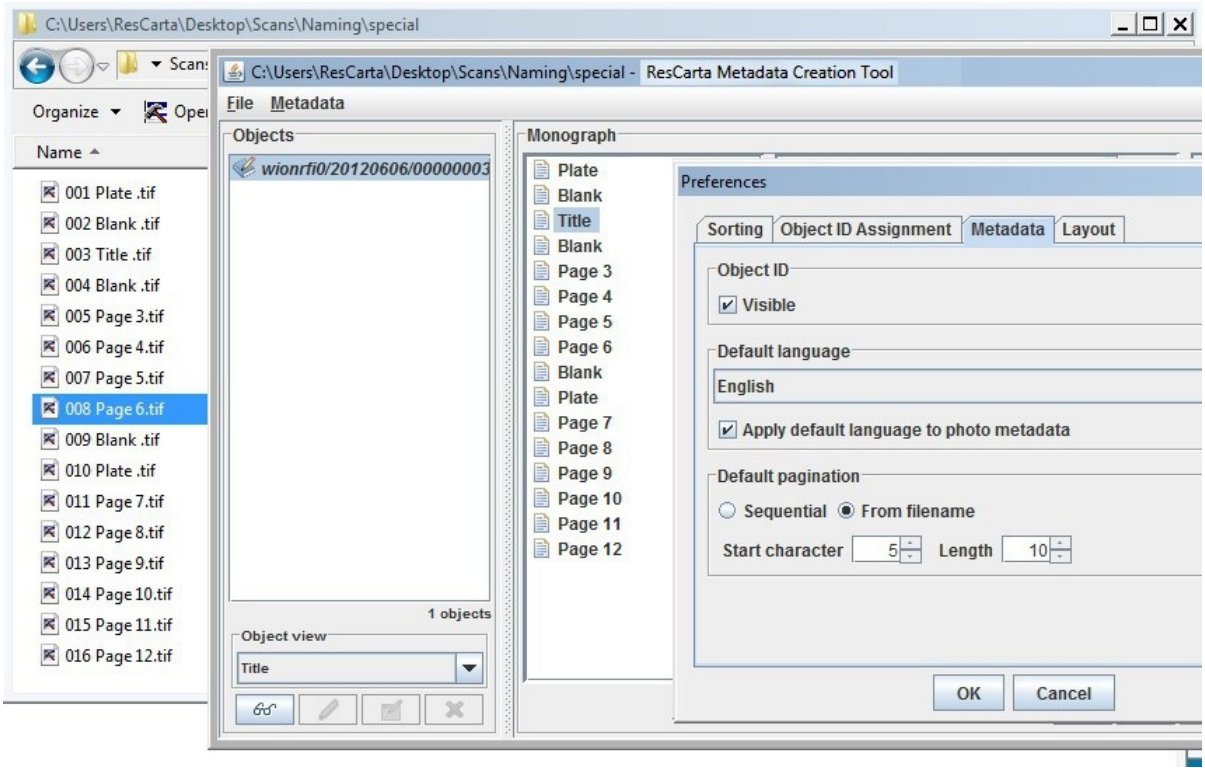


Metadata Tab

This dialog allows the tool to make the Object Identifier visible. If checked the Institutional Identifier, the Aggregator ID and the Root Id will be displayed at the top of the metadata entry form.

It can be used to set the default language to something other than English and to allow the default language to be applied to photographs where the default is none.

The dialog also allows for extraction of page names from file names. The default is to name the pages sequentially. In the case of MOA1 formatted data the page names are found in the final four characters of the file name, so setting the radio button to "From filename", the "Start character" to 5 (fifth character in the filename) and the "Length" to 4 will allow the page names to be extracted from the filename. This will also work with other specialized image file names (e.g. 001Page 1.tif, 002Page 2.tif, 003Page 3.tif) by setting start at 5 and Length at 10 (The Length can exceed an anticipated end, since the file name will be stopped by the extension. See the screen shot below.



Layout Tab

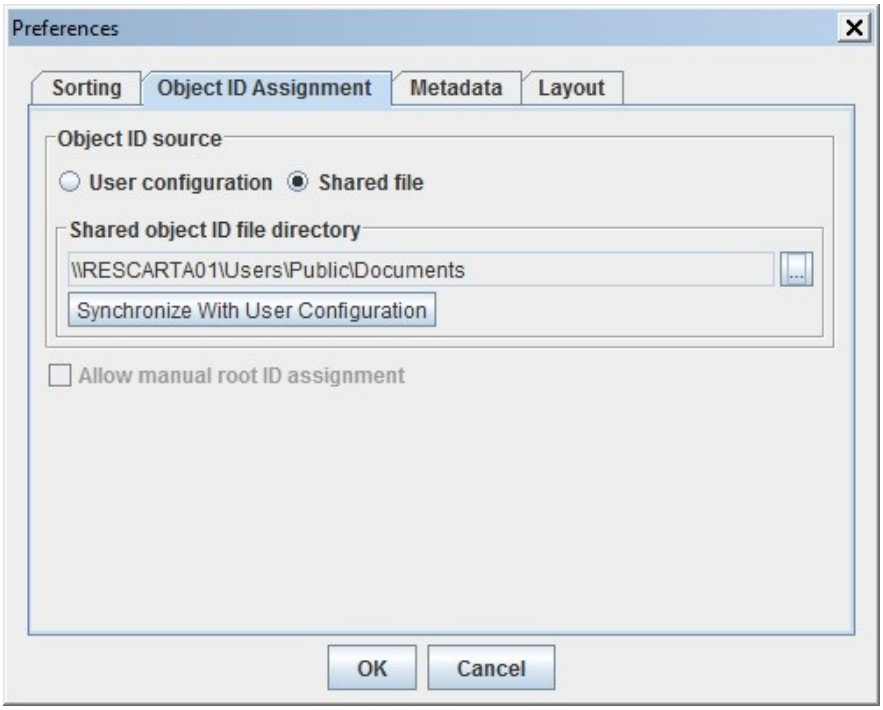
This dialog allows you to change the position of the image viewer and page metadata pane in relation to the metadata entry form. Depending on your screen aspect ratio you might have more room for its placement on top, bottom, left hand or right hand side. Click the radio button to select your preference.

Object ID Assignment Tab

This dialog is used to set the location and flexibility of the RcObjId.xml file. This file can be shared among users of the ResCarta Metadata Creation Tool to coordinate Institutional Identifiers, Aggregators and Root Identifiers.

Object ID Source: This will set the location of the file where your Institutional Identifiers are stored. The file can be located in the home directory of a local user or a network shared location. If you have started using ResCarta tools alone and want to share the information with a team, select **Shared File**, enter or browse for a network location where you and your team have write permission and then press **Synchronize with User Configuration** this will allow you to share your existing Institution Identifiers and coordinate your team.

Allow manual root ID assignment: This checkbox is only available if you are in single "User configuration" mode. By checking this box you will be allowed to edit the Root ID field when beginning to assign new metadata to an object. You will be responsible for creating unique identifiers for each object within an Aggregator. If you select "Shared file" mode this check box is not allowed since the ResCarta Metadata Creation Tool must keep track of Aggregators and Root Identifiers and keep them unique among several users.

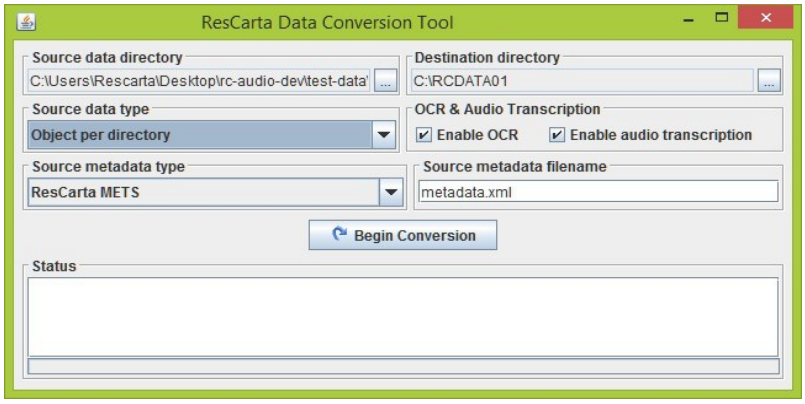


ResCarta Data Conversion Tool

Use the **ResCarta Data Conversion Tool** to convert your TIFF, JPEG, PDF (image only), PDF (image and text) or Wav files into *ResCarta archive data format*. Any new, digitized object that you create for inclusion in your ResCarta archive needs to be converted into this open archive format in order to be recognized by subsequent tools.

To convert your images into ResCarta format:

1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **2.-ResCarta Data Conversion Tool**.



   Note: The version number of the **ResCarta Data Conversion Tool** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

   The **ResCarta Data Conversion Tool** appears.

2. Click the button to the right of the **Source data directory** text field to specify the directory in which your source documents are located. ...

   The Select directory dialog appears.

3. Enter or browse to the **Source data directory** (the directory in which the documents you want to convert are stored) and click **Select**. Select

   Note: Because the ResCarta Standard requires that all images representing a particular document are stored in the same directory and all directories representing documents are stored under the same "parent" directory, *enter or browse to that parent directory* to work with any or all of the documents included within it.

4. Click the button to the right of the **Destination directory** text field to specify the directory into which your converted documents will be placed. ...

   The Select directory dialog appears.

5. Enter or browse to the **Destination directory** (the directory into which the converted documents will be placed) and click **Select** directory. Select

Important! The destination directory can be located wherever you specify, if it does not exist at the time you begin the conversion process it will be created, it **must** be named **RCDATA01 (upper case)**.

Safety Note: As a good practice do *not* use your ResCarta archive (the intended final location of your data) as your destination directory; use an intermediate location during the conversion process, and then transfer converted data from this intermediate location to your ResCarta archive.

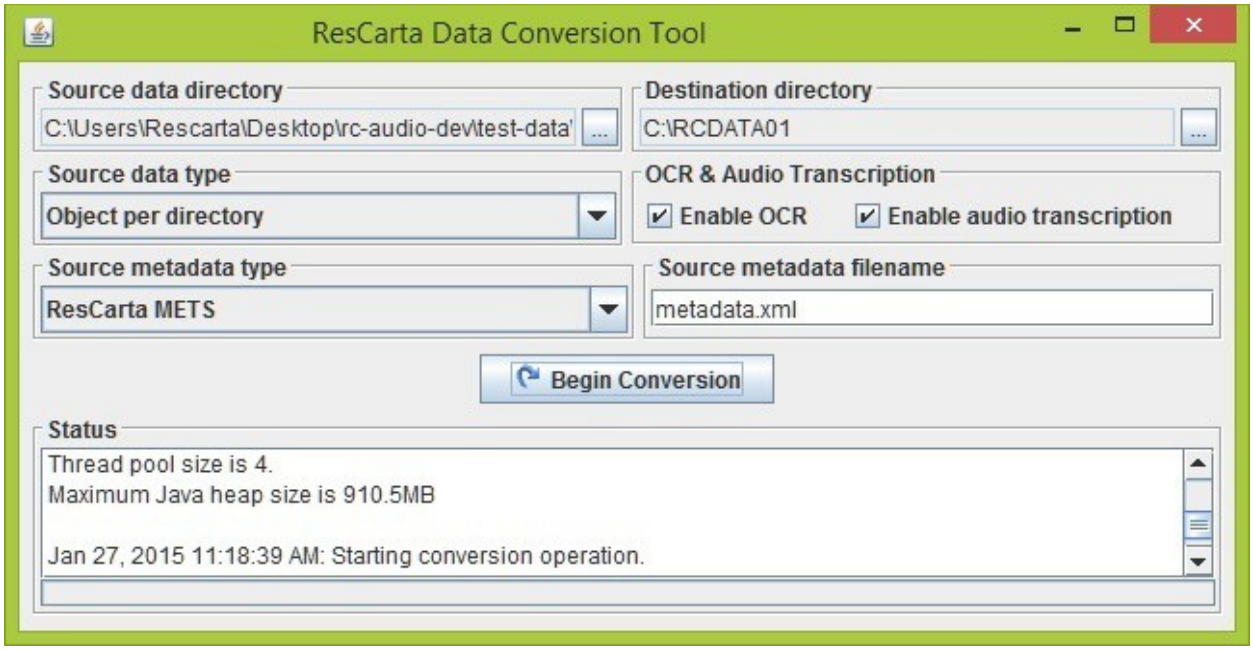
6. Select the Source data type as **Document per directory** from the pull down and select the **Source metadata type** as **ResCarta Mets**.

   Note: The **Source metadata type** for Documents per directory is currently restricted to **ResCarta METS**. Future versions of the **ResCarta Data Conversion Tool** will include support for other source metadata types.

7. Enter the name of the **Source metadata filename** as **metadata.xml**, and click **Begin conversion**.

   Note: If you created metadata using the **ResCarta Metadata Creation Tool** (see page 6), the source metadata filename is **metadata.xml**; if the source metadata filename you specify does not exist in a document directory, *the conversion process will fail*.

   The progress of the conversion process is detailed in the **Status** (lower) section of the **ResCarta Data Conversion Tool**, and you will be notified when the process is complete.



   Note: Depending on the number and size of the documents in your source data directory, the conversion process may take some time; for larger data sources, or when running OCR conversions, you may want to run your conversion process overnight. OCR conversion is performed using **Tesseract Open Source OCR Engine.** This can be enabled for conversion by checking the "Enable OCR" checkbox in the main dialog. Converted images are saved in archive TIFF format with metadata preserved in the file in MODS (Metadata Object Description Schema) format and externally in ALTO (Analyzed Layout and Text Object) format. For more information about MODS/ALTO, refer to the website at the Library of Congress: http://www.loc.gov/standards/alto or http://www.loc.gov/mods

8. When the conversion is completed, click the **ResCarta Data Conversion Tool**'s *close window* button. 

   Note: The **ResCarta Data Conversion Tool** creates the following directory structure under your **RCDATA01** (*destination directory*, as specified in step 5): *institution identifier/aggregator/root identifier*. These three terms, which comprise the *object identifier* for a particular document, are defined when you use the **ResCarta Metadata Creation Tool** (see step 10, page 8). For example, if your **Id** is **brookhav**, your **Aggregator** is **00000001**, and your **Root id** for a particular document is **00000026**, the following directory structure appears for that document under your **RCDATA01** destination directory: **brookhav/00000001/00000026**.

Data Conversion tool "Object by file" mode:

The Object by file mode of the Data conversion tool allows for the conversion of photographs, multipage PDF files or audio files into ResCarta objects. It can convert multipage NORMAL, Image and text and Image only PDFs. It converts mono or stereo 24-bit 96Khz and 16-bit 44.1Khz formatted Wav files.

NORMAL PDF files contain fonted text and do not contain a complete raster image of the document page. The ResCarta tool will need the assistance of GHOSTSCRIPT to convert the fonted text data into TIFF raster format. This process used to be known as a RIP ("Raster Image Processor"), which prepared Postscript files for printing on raster laser printers. If you unchecked GHOSTSCRIPT during the installation you can download GHOSTSCRIPT from http://www.ghostscript.com/download/gsdnld.html

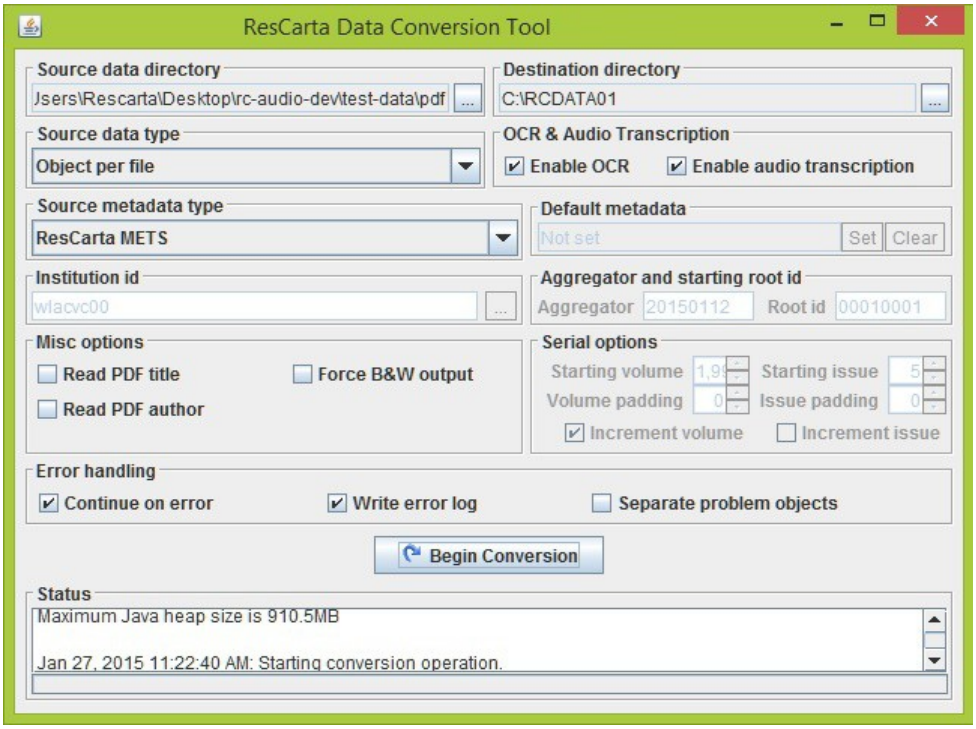
Proceed with the first five steps from the previous section. Then

1. Select **Source data type** as "**Object per file**" from the pulldown.
2. Now you have the **option** of using individual metadata.xml files created earlier using the Metadata Creation tool **OR** apply a set of standard metadata to all the files in the directory by selecting the Source metadata type of "**Apply default metadata to all documents**".
3. **If using** "Apply default" you must set the default metadata to be used. Enter the **Institution Id** for these documents. Enter the starting **Aggregator** and **Root Id** (see Metadata Creation tool on page 8 for more on this.)
4. Set the Default metadata by pressing the **Set** button. You will be asked to **Select the metadata type** (Monograph, Serial, Audio, etc.) and then press the **Ok** button. The Set default metadata dialog will open (see page 9) Enter the items that will be applied to **ALL** the documents in this directory. Then press the **Finish** button at the bottom of the dialog. You may also use a saved metadata template at this time from the **Metadata/ Apply Template** top menu and navigating to to the desired template file.
5. If you would like the tool to extract existing metadata from the PDF files use the **PDF Options** check boxes.
   - Read PDF **title** will extract the title from the existing PDF and if not checked or not found will replace the title with the **file name** from the document.
   - Read PDF **author** will extract the author information if it exists else it will use the author information entered in the **Set default metadata dialog**.
   - Force **B&W Ouput** will create only bitone output from PDF files even if they contain color or grayscale
   - Enable **OCR** will engage the Tesseract OCR engine to create text searchable objects.
   - Enable **audio transcription** will run the automatic transcription generator for audio files.

- **Serial Options** will be grayed out for Monographs. If you are converting a number of Serials you can use the spinners to set the **Starting volume** number, **Starting issue** number and the length of the padding on the Volume and Issue strings. You may also choose to increment the volume or the issue per each file encountered.

6. Error handling can be controlled from the three **Error handling** check boxes.
   - **Continue on error** does just that. The program will continue when it encounters an error.
   - **Write error log** check box will write a log of any files it has problems with or errors encountered during conversion.
   - **Separate problem documents** check box will create a new directory below the existing working directory and move any problem files into that directory.

Press the Begin conversion button when you have completed your selections. The program will show the status of the conversion in the lower pane. You may copy this status information using the operating system's copy/paste key combinations.



Note: see the appendix for details about options to the data conversion process.

ResCarta Textual Metadata Editor

Use the **ResCarta Textual Metadata Editor** to add, edit, or delete *textual metadata* (searchable words that are stored with document images) to ResCarta-formatted documents. Any documents that were scanned without OCR (Optical Character Recognition) can have textual metadata added to them with the use of the Textual Metadata Editor. You may also want to add searchable text areas to photographs, text to image only pages in documents, or to edit existing textual metadata in OCR-scanned pages.

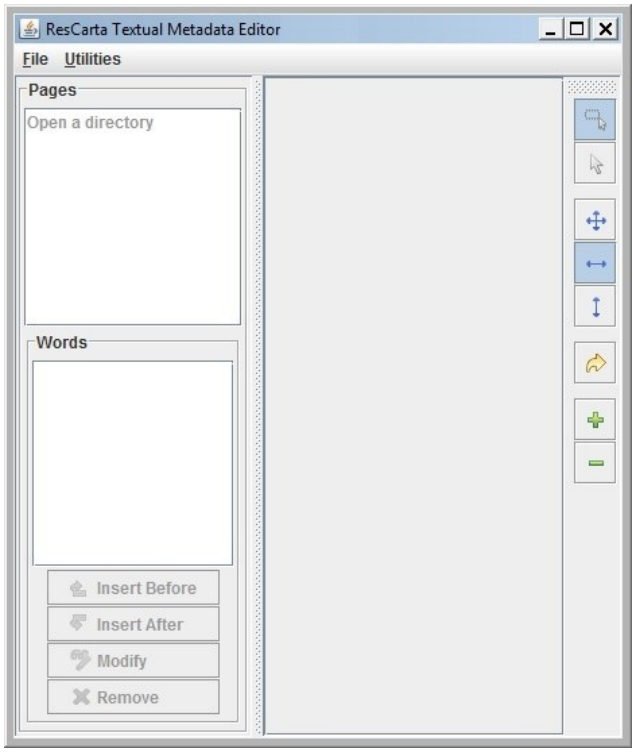
Note: You must use the **ResCarta Data Conversion Tool** (see page 22) *before* using the **ResCarta Textual Metadata Editor**. The Textural Metadata Editor will not work on raw unformatted TIF images that are not in a ResCarta directory structure.

To add or edit textual metadata:

1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **3.-ResCarta Textual Metadata Editor**.

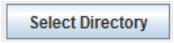
   Note: The version number of the **ResCarta Textual Metadata Editor** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

   The **ResCarta Textual Metadata Editor** appears.

2. 

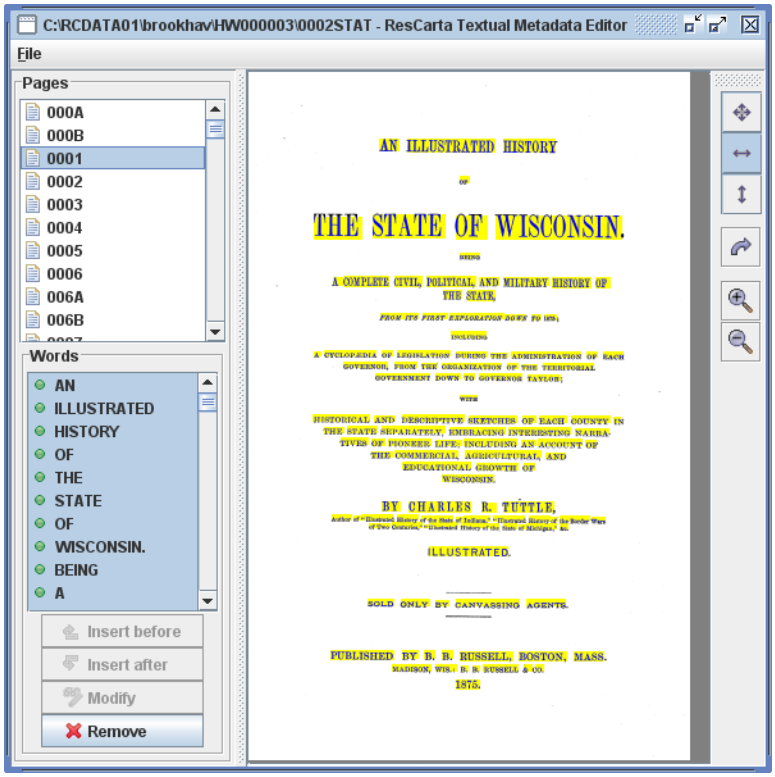
3. From the **File** menu, select **Open ResCarta object directory**.

   The Select ResCarta object directory dialog appears.

4. Enter or browse to the *ResCarta object directory* (document directory; the lowest level of directory structure in documents that have been converted to ResCarta format) in which you want to add or edit textual metadata and click **Select directory**. 

Note: If you used the **ResCarta Data Conversion Tool** to convert your data in the ResCarta format, the following directory structure exists under your **RCDATA01** (*destination directory*, as specified in step 5, page 22 in the **ResCarta Data Conversion Tool** section): *institution identifier/aggregator/root identifier*. These three terms, which comprise the *object identifier* for a particular document, are defined when you use the **ResCarta Metadata Creation Tool** (see step 10, page 8); they also define your ResCarta object directories.

5. The **ResCarta Textual Metadata Editor** now displays a list of **Pages** in the document within the ResCarta object directory you selected (in the upper left pane) and the existing textual metadata (**Words**) for the current page (in the lower left pane). The rest of the tool displays the current page (in the right pane).
6. To select a page, click its number in the **Pages** list. Any textual metadata (**Words**) that exists for that page is listed and selected (In the words pane). The page display (in the right pane) highlights any of the **Words** for which coordinate location values exist, as shown in the following example:

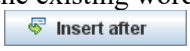


Note: The page display area includes some of the same viewing options that are available in the **ResCarta Metadata Creation Tool**'s viewer (for more information, see page 8).

7. To add textual metadata, use the *insert* button that indicates the location you want to place your new word in the list of **Words**:

Tip: If the order of words in the textual metadata (and in the **Words** display list) is not important to you, it is immaterial which *insert* button you use; the order of the words is not relevant to their coordinate positions, if any are specified at all.

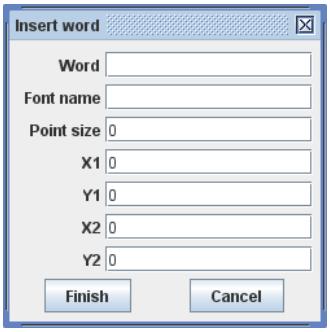
- To add a word *before* another word, select the existing word (the word before which you want to insert a new word) and click **Insert before**. 

- To add a word *after* another word, select the existing word (the word after which you want to insert a new word) and click **Insert after**. 

- To add a word on a page that doesn't yet include any words, click **Insert**. 

Note: If the current page doesn't yet include any words, the **Insert** button replaces the **Insert before** button; if more than one word is selected, **Insert before** and **Insert after** are not available.

The Insert word dialog appears. Enter the information for the word you're inserting:



Tip: Only the **Word** and a **Font name** are *required* to store your word as textual metadata associated with the current page; all other information is optional. However, words for which **X1**, **Y1**, **X2**, and **Y2** values are not entered cannot be highlighted during searches of your collections.

Enter the **Word** you're inserting.

Enter the **Font name** that was used to render the word on the page.

Tip: If you don't know this information exactly, enter the font name that best approximates the displayed font (e.g. Helvetica, Times-Roman).

Enter the **Point size** of the word on the page.

Tip: The Point size is currently reserved for future use; if you don't know this information, enter the point size that best approximates the displayed size.

Note: **X1**, **Y1**, **X2**, and **Y2** are measured in pixels at 72 DPI, with the lower left corner of the *page* as the origin. You can use the  **Highlight Editing** tool to drag a box to populate the coordinates or

Enter the **X1** coordinate of the word (lower left X coordinate).

Enter the **Y1** coordinate of the word (lower left Y coordinate).

Enter the **X2** coordinate of the word (upper right X coordinate).

Enter the **Y2** coordinate of the word (upper right Y coordinate).

8. To edit textual metadata for an existing word, select the word you want to edit and click **Modify**. 

9. *Note*: If the current page doesn't yet include any words or more than one word is selected, **Modify** is not available.
10. The Modify word dialog appears. Enter the information for the word you're modifying (refer to the preceding step for more details on the information that's required in this dialog).
11. To remove a word from the current page, select the **Word** you want to remove and click **Remove**. 

The Confirm word removal window appears. Click **Yes** to remove the word from the current page; click **No** to cancel the word removal process.
12. When you're finished adding and editing textual metadata, click the **ResCarta Textual Metadata Editor**'s *close window* button.  The confirm exit window appears. Click **OK** to exit; click **Cancel** to return to the **ResCarta Textual Metadata Editor**.

ResCarta Audio Transcription Editor

Use the **ResCarta Audio Transcription Editor** to add, edit, or delete *textual metadata* (searchable words that are stored with audio files) to ResCarta-formatted objects. Any ResCarta formatted Broadcast Wave audio files can have textual metadata added to them with the use of the **ResCarta Audio Transcription Editor**. You may also want to edit existing textual metadata in ResCarta formatted Broadcast Wave audio files.

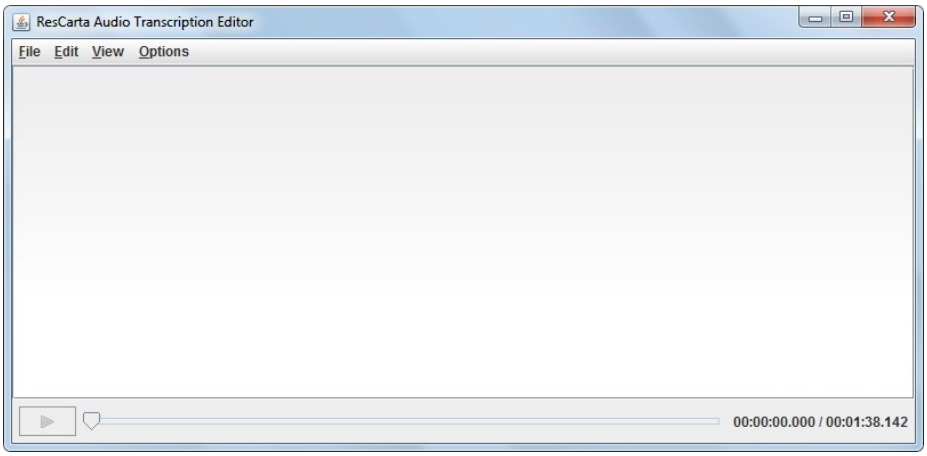
Note: You must use the **ResCarta Data Conversion Tool** (see page 22) *before* using the **ResCarta Audio Transcription Editor**. The editor will not function on raw WAV files.

To add or edit textual metadata:

1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **4.- ResCarta Audio Transcription Editor**.

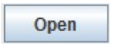
Note: The version number of the **ResCarta Audio Transcription Editor** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

The **ResCarta Audio Transcription Editor** appears.

2. 

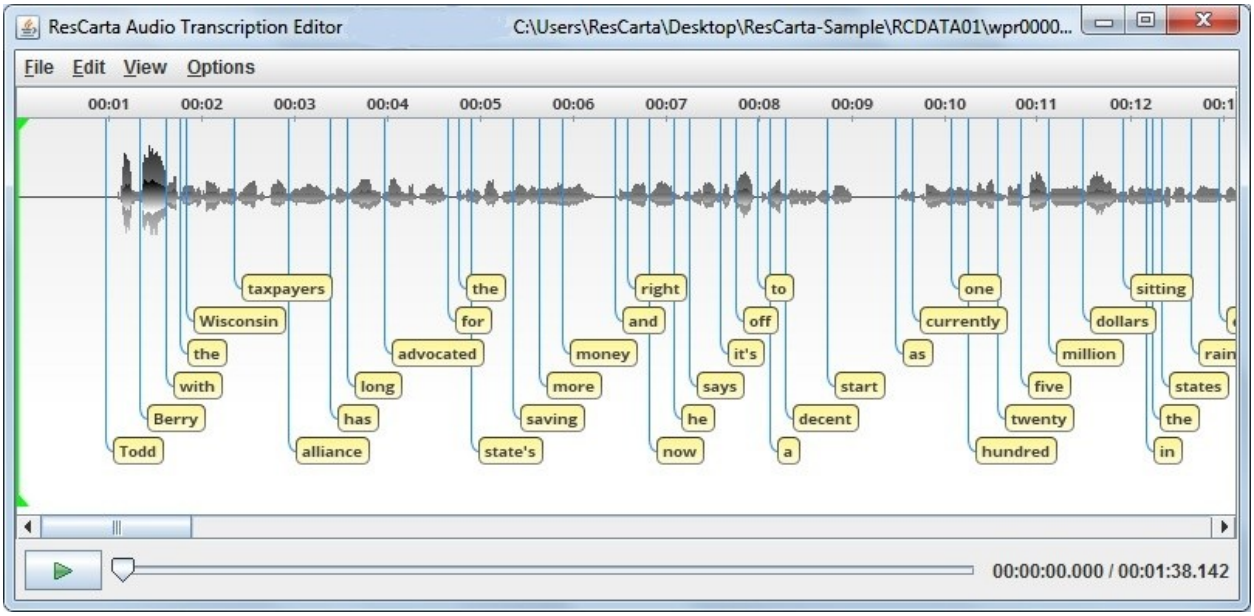
3. From the **File** menu, select **Open**.

   The Select ResCarta Audio object directory dialog appears.

4. Enter or browse to the *ResCarta Audio object directory* (document directory; the lowest level of directory structure in objects that have been converted to ResCarta format) in which you want to add or edit textual metadata and click **Open**. 

   Note: If you used the **ResCarta Data Conversion Tool** to convert your data in the ResCarta format, the following directory structure exists under your **RCDATA01** (*destination directory*, as specified in step 5, page 22 in the **ResCarta Data Conversion Tool** section): *institution identifier*/*aggregator*/*root identifier*. These three terms, which comprise the *object identifier* for a particular document, are defined when you use the **ResCarta Metadata Creation Tool** (see step 10, page 8); they also define your ResCarta object directories.

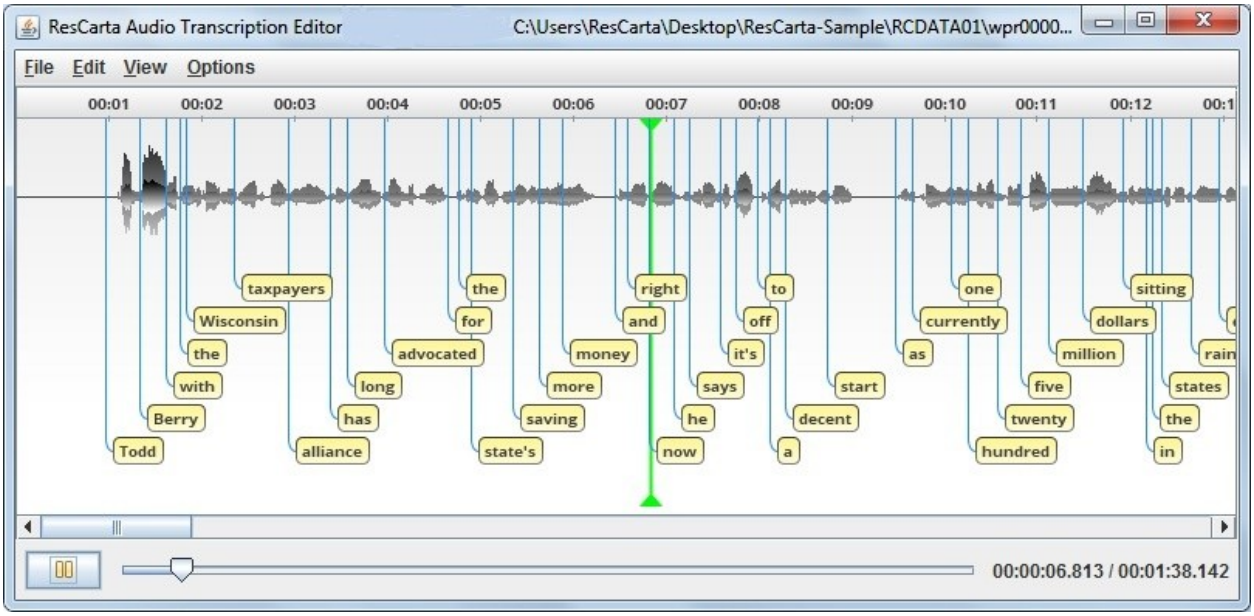
5. The **ResCarta Audio Transcription Editor** now displays a waveform icon with text boxes located within the audio file.



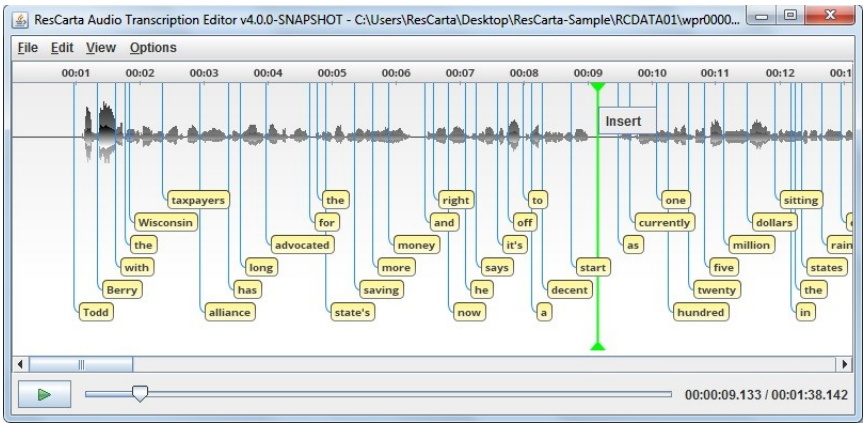
6. Click the play audio button in the lower left hand corner  to review the existing text while the audio playback proceeds. A green location bar will appear showing the location in the audio file.

   Note: The text display will scroll to the right as the audio timeline moves beyond the size of the window.

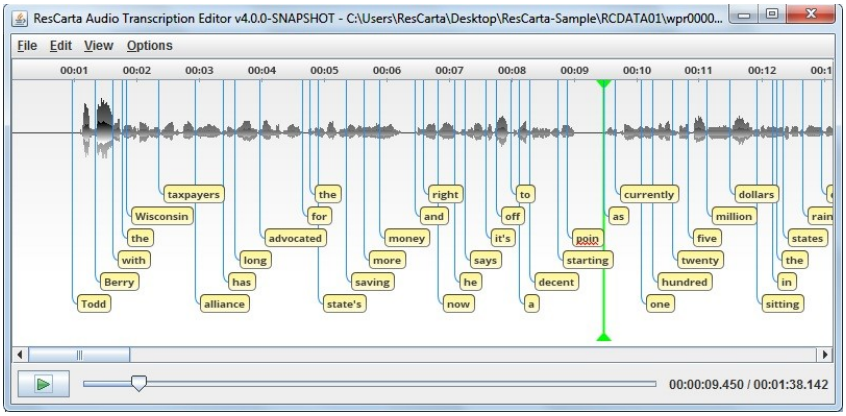
7. To stop playback at any time click the pause audio playback button  .



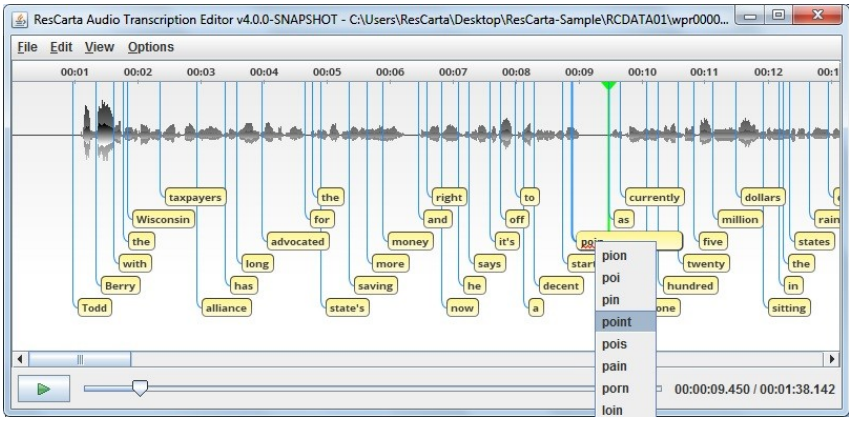
8. Once playback has been paused you can begin to edit the existing text or add more text to the audio file. To correct a word click on the **word box** with your mouse. The word will be selected and you can now type a new word into the box or click into the word with the mouse to add/edit characters within the word.



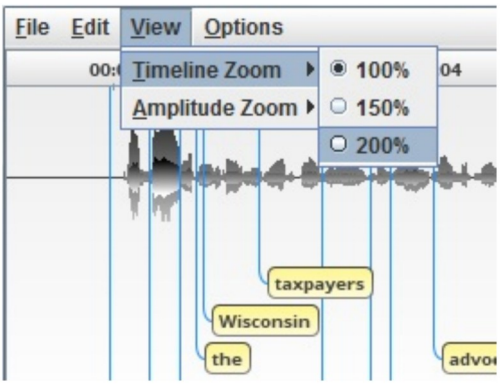
9. To add textual metadata, click on the waveform between existing word locations, the green location bar will indicate the start of the text. Right click to get the *insert* menu. Click the Insert menu and a new word box will appear. Enter the desired text into the word box. As you type the spell checker will test the text being entered.

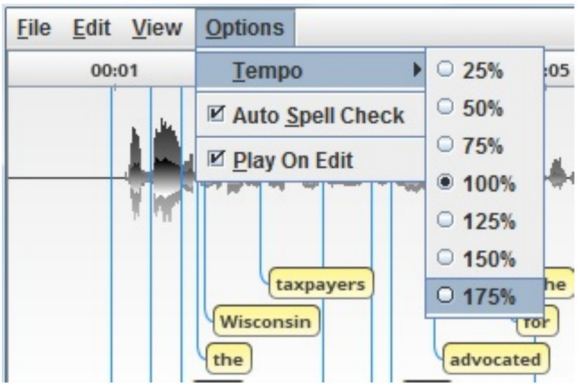


10. To correct a word using the spell checker, click your mouse on a red underlined word. Then right click to get a list of possible corrections. Click the correct word with the either mouse button.



11. Once you have focus on a **word box**, you can navigate through the word boxes by using the up and down arrows on your keyboard. Once inside of a word box, you can use the left and right arrow keys on your keyboard to navigate through the characters within the box.
12. You can vary the speed of playback by selecting the Options menu then Tempo. The Options menu also allows you to disable spell checking and playback on edit. The View menu allows you to adjust the spread of the Timeline and the size of the wave form.



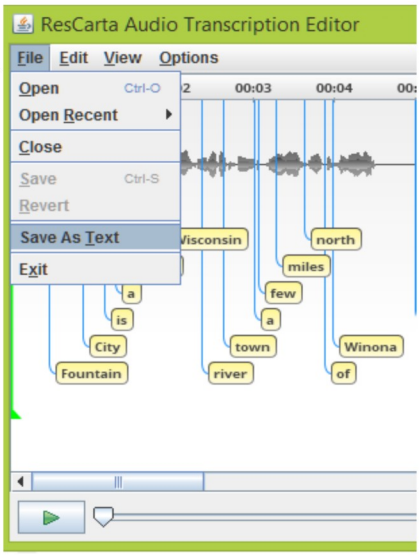


13. The edit menu has an undo and a redo function which can remove the last changes to the file. These can also be accomplished by use of the keyboard combinations of Ctrl-Z and Ctrl-Y

Tip: It is best to save your work often while editing the textual content of your audio files.

14. If you have used the ResCarta Audio Transcription Editor to modify audio files in your archive, the program will keep a listing of recently edited files. You can open a previously edited file by using the File/Open Recent menu item.

15. You can save the transcribed text to a file by using the File/Save As Text menu item.



ResCarta Collections Manager

Use the **ResCarta Collections Manager** to manage your ResCarta collections. Any change you want to make to your collections, such as adding or removing documents, is managed through this tool.

Important! You should copy your ResCarta data from its intermediate location to your ResCarta archive (its intended final location) *before* using the **ResCarta Collections Manager**. Use a file manager (such as Windows Explorer) to locate your intermediate **RCDATA01** folder, and copy the directory or directories immediately beneath this folder (your *institution identifier* directory, as specified in the **ResCarta Metadata Creation Tool** section, step 10, page 8; if you've used more than one institution identifier, you will have more than one directory to copy) to the same directory level of the computer on which your ResCartaWeb archive is stored (immediately beneath the **RCDATA01** folder).

Note: After making any changes to your ResCarta collections, *objects* (ResCarta-formatted documents), or metadata, the **ResCarta Indexer** must be run to update the index (see page 41).

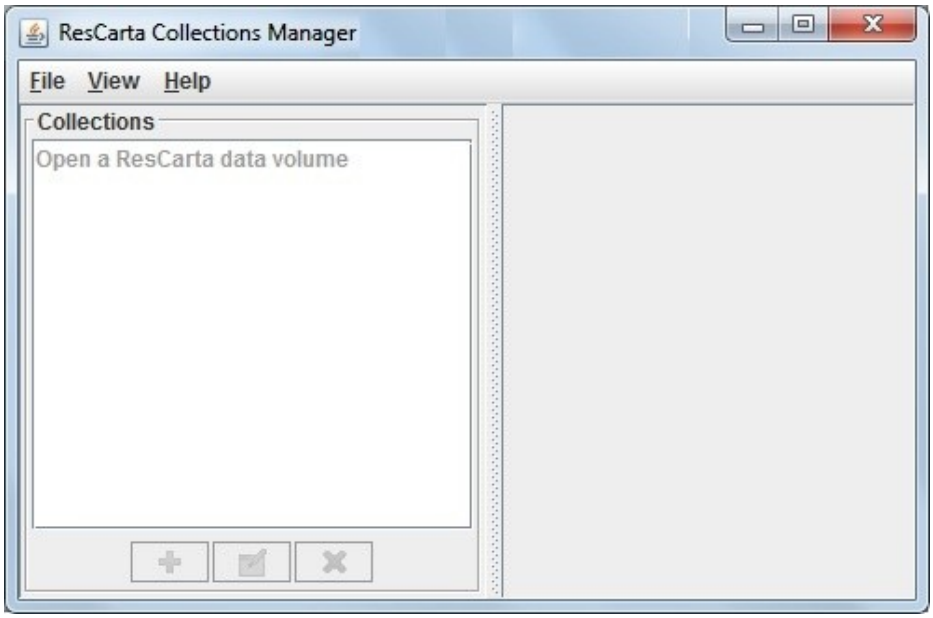
To manage your ResCarta collections:

Note: The following information is also accessible from within the **ResCarta Collections Manager**; from the **Help** menu, select **Help contents**.

1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **5.-ResCarta Collections Manager**.

   Note: The version number of the **ResCarta Collections Manager** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

   The **ResCarta Collections Manager** appears.



2. To open a ResCarta Data Directory:

   Note: The ResCarta data directory will be named **RCDATA01(uppercase)**.

   - From the **File** menu, select **Open ResCarta volume**.

   The Select directory dialog appears.

   - Enter the path to the directory that contains the data you want to work with in the **File Name** box or browse to the directory and then click **Open**. Open

   Note: If you have a large number of ResCarta documents in the directory you're opening, you may notice a delay as the documents are accessed.

   - After the ResCarta Data Directory you selected opens, a dialog appears to inform you that it opened successfully and details the number of ResCarta objects that were found there. Click **OK** to close this dialog.

3. The **ResCarta Collections Manager** displays **Collections** in the volume you selected (in the left pane) and information about the current collection (in the right pane).

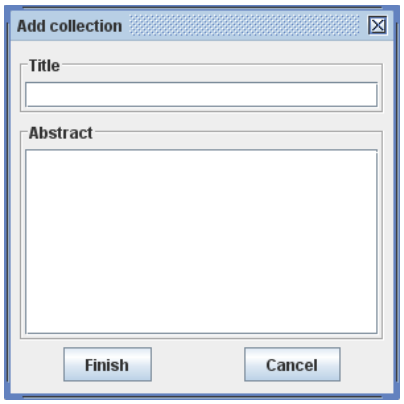
   Note: **Collections** are indicated with a stacked book icon (in the left pane). **Objects** (documents) in collections are indicated with a book icon (in the right pane), with the total number of documents in the current collection displayed at the bottom of the **Objects** list.

   Tip: To copy an *object identifier* (see step 10, page 8) or the complete path to an object, move your mouse pointer over the object for which you want to copy information and right-click; from the menu that appears, select **Copy object id to clipboard** or **Copy path to clipboard**, respectively.)

4. **To create a collection**:

   Click the **Add** button in the **Collections** pane on the far left of the **ResCarta Collections Manager**.

   The Add collection dialog appears.



   Enter the **Collection title** you want to use, enter an abstract of the collection and then click the **Finish** button.

   Note: The collection title can contain letters, numbers, and/or spaces.

   Your new collection title appears in the **Collections** list and the total number of objects in the current collection is displayed at the bottom of the list.

5. To add objects to a collection:

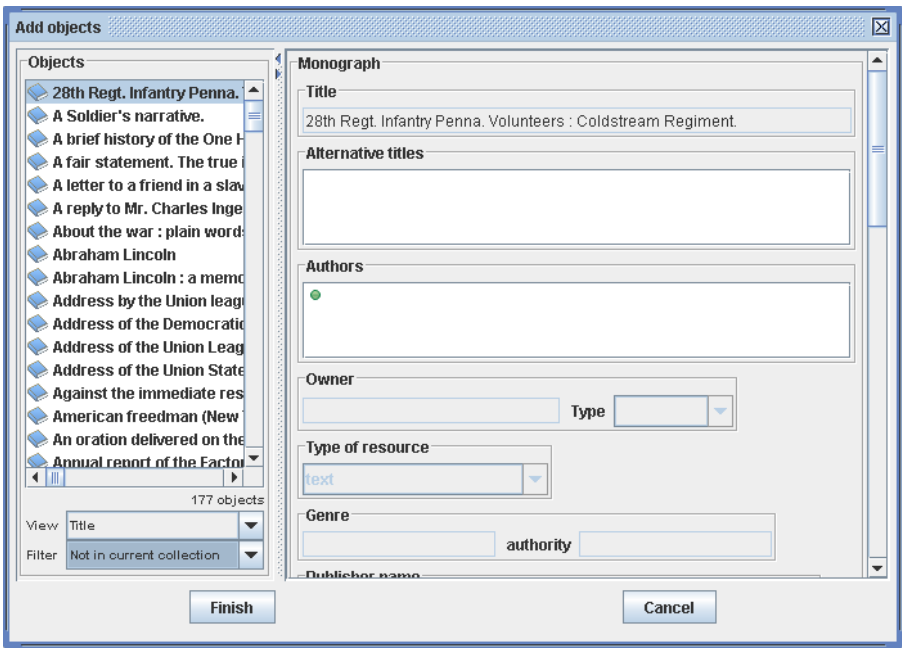
   Select the collection you want to add objects to in the **Collections** pane (on the left side of the **ResCarta Collections Manager**).

   Objects in the selected collection are displayed on the **Collection content** tab (in the right pane); the total number of objects in the current collection is displayed at the bottom of the **Objects** list.

   Tip: Use the **View** drop-down at the bottom of the **Objects** section to display objects by **Title** or **Id** (*object identifier*; see step 10, page 8).

   - On the **Collection content** tab, click the **Add** button under the list of **Objects** that are currently included in the collection.

   The Add Objects dialog appears, listing all available data in the current ResCarta Data Directory); the total number of objects *with respect to the current* ***Filter*** (see the following tip) is displayed at the bottom of the **Objects** list.



   Note: Metadata for the currently selected object is displayed (in the right pane of the Add Objects dialog) for reference; this information can be edited *after* the object is added to a collection (see the following step). The page display area (at the bottom of the right pane of the Add Objects dialog) includes some of the same viewing options that are available in the **ResCarta Metadata Creation Tool**'s viewer (for more information, see page 8).

   Tip: Use the **View** drop-down at the bottom of the **Objects** section to display those documents by **Title** or **Id** (*object identifier*; see step 10, page 8). Use the **Filter** drop-down at the bottom of the **Objects** section to display **All** (all objects in the your ResCarta volume), **Not in current collection** (only objects in your ResCarta volume that are not included in the current collection, as shown in the preceding example), or **Not in any collection** (only objects in your ResCarta volume that are not included in any collection).

- Select the object or objects you want to place in the collection.
- To select a single object, click it.
- To select multiple objects that appear adjacent to one another in the list, press and hold the **Shift** key as you click the *first* and *last* of the adjacent objects you want to place in the collection; the rest will be selected automatically.
- To select multiple objects that do not appear adjacent to one another in the list, press and hold the **Ctrl** key as you select the objects you want to place in the collection; you must click each individual object you want to select.
- Click **Finish** to add the selected object or objects to the collection.

Note: You may add the same object to more than one collection; the metadata will be maintained independently.

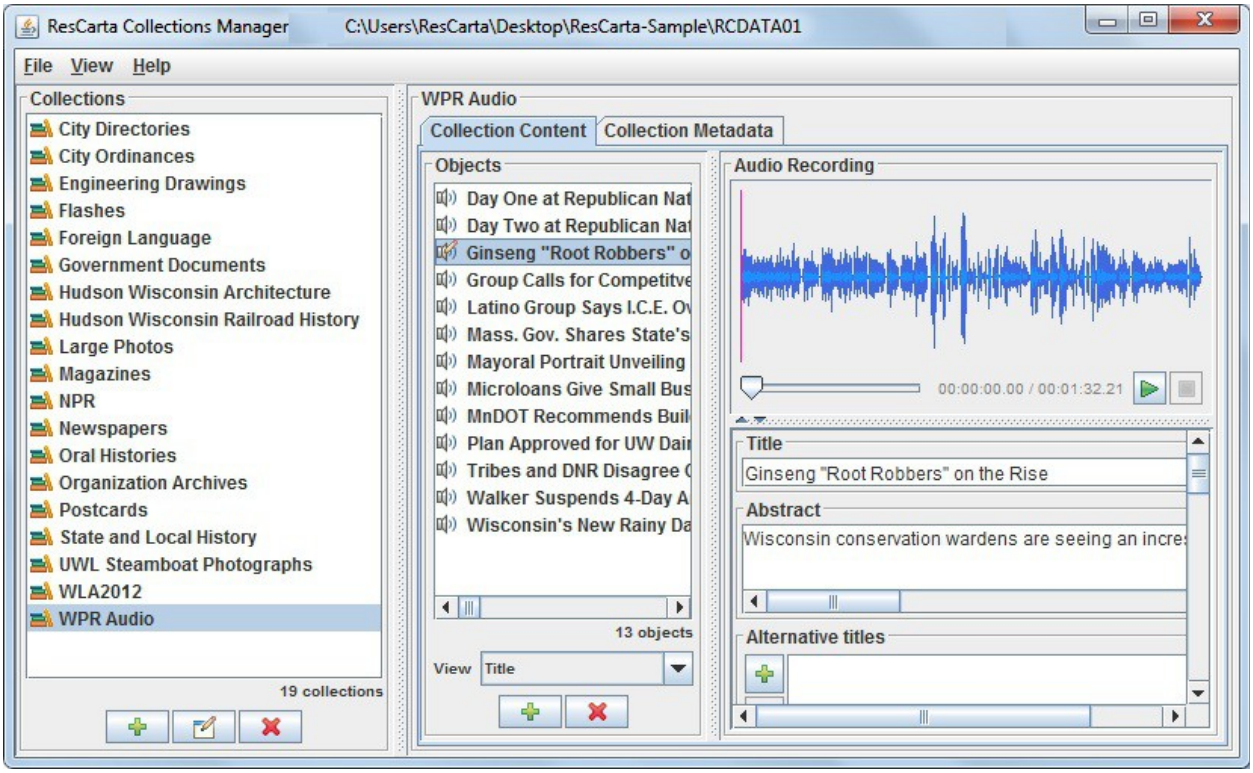
6. To remove an object from a collection:
    - Select the collection you want to remove an object from in the **Collections** pane (in the left pane).
    - Objects in the selected collection are displayed (in the right pane, in the **Collection content** tab).
    - On the **Collection content** tab, select the object or objects you want to remove from the list of **Objects** that are currently included in the collection and click the  **Remove** button under the list.
    - To select a single object, click it.
    - To select multiple objects that appear adjacent to one another in the list, press and hold the **Shift** key as you click the *first* and *last* of the adjacent objects you want to place in the collection; the rest will be selected automatically.
    - To select multiple objects that do not appear adjacent to one another in the list, press and hold the **Ctrl** key as you select the objects you want to place in the collection; you must click each individual object you want to select.
    - The Confirm object removal dialog appears to confirm the object you want to remove, and the collection you want to remove it from.
    - Click **Yes** to remove the object from the collection. Click **No** to close this dialog without removing the object from the collection.

Note: Removing an object from a collection does *not* remove it from its original location in the ResCarta Data Directory or delete any data in the archive.

7. To modify object metadata:

   Note: Changes made to metadata in one collection do not occur in that same object in other collections. **Size width**, **Size height**, **Issues**, and **Pages** cannot be modified using the **ResCarta Collections Manager**.

   - Click the object for which you want to edit metadata.



   Note: Use the vertical scroll bar on the right to navigate to the fields you want to modify.

   - To modify **Title**, **Volume** (for serial monograph or serial documents only), **Owner**, **Genre**, **Genre authority**, **Publisher name**, **Place of publication**, **Publication date**, **Capture date**, **Language**, and **Note** metadata, select the field you want to change and make the changes directly to the field in the Modify Objects dialog. The title of the field will turn green to note editing.

   - To modify **Owner Type**, **Type of resource**, **Language authority**, and **Language type**, choose from the options in the drop-down list in the Modify Objects dialog.

   - To modify **Alternative titles**, **Names**(Authors, etc.) and **Subjects** metadata, use the **Add**, **Modify**, and **Remove** buttons located at the left hand of the field you're modifying.

8. To add information:

   Click the **Add** button under the field to which you want to add information. 

   The Add dialog appears for the metadata you selected.

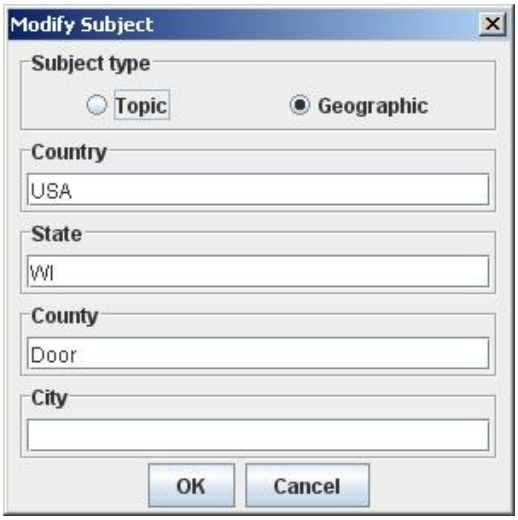
   Enter the information you want to add to the field and click **Finish**. To close this dialog without saving your changes, click **Cancel**.

   Note: When you're adding **Subjects** metadata, you must first select **Subject type** (**Topic** or **Geographic**). For **Topic** type metadata, enter **Authority** and **Topic** information. For **Geographic** type metadata, enter **Country**, **State**, **County** and **City** information.

9. To edit existing information:

   If multiple lines of information exist for the metadata you want to change, select the specific line you want to modify and then click the **Modify** button under the field. 

   The Modify dialog appears for the metadata you selected.



Make the changes to the information that you want to include in the field and click **Finish**. To close this dialog without saving your changes, click **Cancel**.

Note: When you're changing **Subjects** metadata, you may need to change **Subject type** (**Topic** or **Geographic**). For **Topic** type metadata, you can change **Authority** and **Topic** information. For **Geographic** type metadata, you can change **Country**, **State**, **County and City** information.

o

10. To remove information:

    If multiple lines of information exist for the metadata for which you want to remove information, select the specific line you want to remove and then click the **Remove** button to the left of the field.

    The Confirm removal dialog appears for the metadata you selected, to confirm the information you want to remove. Click **Yes** to remove the information. Click **No** to close this dialog without removing the information.

    - .

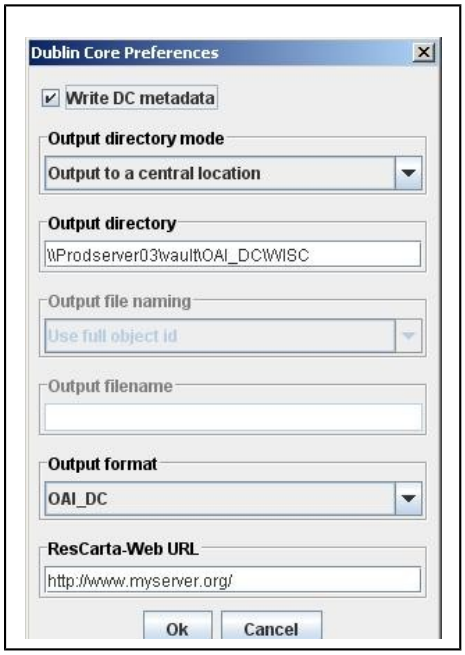
11. To save your changes to your ResCarta collections, from the **File** menu, select **Save collections data**.

    Tip: When you're making a large number of changes, save your changes periodically rather than waiting to save your changes when you exit the **ResCarta Collections Manager**.

12. When you're done managing your ResCarta collections, click the **ResCarta Collections Manager**'s *close window* button. The confirm exit window appears. Click **OK** to exit; click **Cancel** to return to the **ResCarta Collections Manager**.

    Note: If there are changes to your collections that you haven't yet saved, you will be prompted to save those changes before exiting the **ResCarta Collections Manager**.

13. **Saving Dublin Core Metadata.** You can save your METS metadata to Dublin Core formatted xml files for use in other systems that require the more generic metadata format. From the **File** menu, select **Dublin Core Preferences**. The Dublin Core Preferences dialog appears.



For Oai_DC

To Create OAI_DC xml files for use in your OAI provider software.

1. Set the output directory to a central location using the pulldown.
2. Enter the path of the output directory.
3. Select OAI_DC as the output format from the output pulldown.
4. Select "Use full object id" from the output file naming pulldown
5. Enter the url of your ResCarta –Web site complete with the trailing slash.

(e.g. http://www.ResCarta.org/ResCarta-Web/)

6. Press the **Ok** button
7. Answer **Yes** to the write dialog. This will create an xml file in Oai_DC format for each object in the archive that has been added to a collection. Objects not in a current collection will not have a corresponding xml file created.

For Dubin Core

To create simple Dublin Core metadata files for each object.

1. Set the output directory either to a central location or to the source object folder. Selecting source object folder will place the Dublin Core xml file within the same folder as the source METS metadata file.
2. Set the output file naming pulldown to either "full object id" or "custom name" from the ouput file naming pulldown. If you select "custom name" you will need to enter a filename into the Output filename box.
3. Set the Output format pulldown to DC.
4. Press the Ok button.
5. Answer **Yes** to the write dialog. This will create an xml file in Duboin Core format for each object in the archive that has been added to a collection. Objects not in a current collection will not have a corresponding xml file created.

ResCarta Indexer

Use the **ResCarta Indexer** to create a Lucene™ index of your ResCarta collections. Any changes you make to your collections (using the **ResCarta Collections Manager**; see page 33) require that a new index be built so that ResCartaWeb can accurately search the collection.

Important! Indexing a collection can be a time-consuming process, particularly in larger collections; the ResCarta Foundation recommends running **ResCarta Indexer** over a weekend.

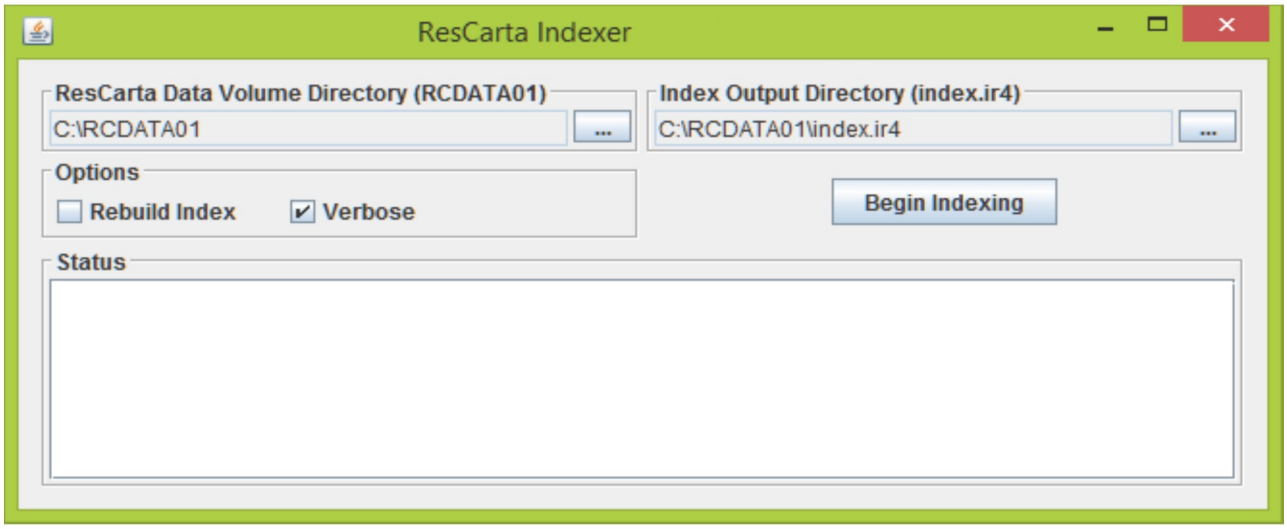
To create Lucene indexes of your collections:

Note: The following information is also accessible from within the **ResCarta Indexer**; from the **Help** menu, select **Help contents**.

1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **6.- ResCarta Indexer**.

   Note: The version number of the **ResCarta Indexer** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

   The **ResCarta Indexer** appears.



   If you're creating an index for a ResCarta collection for the first time, set your **Source data directory** and **Index destination directory** and then click **Begin indexing**. The indexing process will begin and its status will be displayed in the **Status** window.

   If you've made structural changes to your ResCarta collection metadata at the *collection level* (for example, adding or removing a collection) *or* object level (for example, adding or removing a book from a collection), you can also proceed to **Begin indexing** (see step 1 below) as soon as you've set your **Source data directory** (see next step) and **Index destination directory** (see step 3).

   Other changes to the ResCarta collection metadata may require using of indexing **Options** (see step 4).

2. To select the **Source data directory**:
   - Click the button to the right of the **Source data directory** text field. ...
   - The Select directory dialog appears.
   - Enter the path to the directory that contains the data you want to work with in the **File Name** box or browse to the directory and then click **Select directory**. Select directory

   Note: The ResCarta "source data directory" will be named **RCDATA01(uppercase)**.

3. To select the **Index destination directory**:

   Important! ResCarta Web specifications require that indexes be stored in the **index.ir4 (lower case)** subdirectory of the **RCDATA01(upper case)** ResCarta data directory (**/RCDATA01/index.ir4**). If this directory does not yet exist, the index tool will create it.
   - Click the button to the right of the **Index destination directory** text field. ...
   - The Select directory dialog appears.
   - Select **index.ir4** subdirectory:
   - *If the **index.ir4** subdirectory already exists in the **RCDATA01** directory*, enter the path to the **index.ir4** subdirectory or browse to the directory and then click **Select directory**. Select directory

4. To select indexing **Options**:

   There are two options that may be used when indexing: **Force rebuild** and **Verbose**. To use either, or both, select the appropriate check box(es).

   Rebuild index

   When you check **Rebuild index**, the ResCarta Indexer will **delete the existing index** and build a new index, even if t*here have been no changes in the data*. For example, if your index becomes corrupted, you need to use **Rebuild index** to create a new index.

   You will also want to use **Rebuild index** when you have used the **ResCarta Collections Manager** to change metadata for any of the objects that already existed in your collections. For example, if you change the title of a book or the subject, you must run the **ResCarta Indexer** with the **Rebuild index** option checked to create an index that reflects the changes.

   *Note: If you are running the indexer against an archive being served by a ResCarta-Web application, the **TOMCAT server must be STOPPED** in order to perform a complete rebuild of the index. Trying to rebuild an index while it is being served by a web application will result in a corrupted index and a broken website.*

   Verbose

   Using the Verbose check box will output a status for each object being indexed. With the box unchecked there will only be a progress ribbon with a percentage of completion. This does not affect the time to complete the index.

5. The progress of the indexing process is detailed in the **Status** (lower) section of the **ResCarta Indexer**.

6. When indexing is complete, click the **ResCarta Indexer**'s *close window* button.

ResCarta Checksum Verification Tool

Use the **ResCarta Checksum Verification Tool** to verify TIFF *checksums* (computed values of data that are compared to ensure data accuracy) of your ResCarta-formatted data, to ensure the integrity of the documents in your ResCarta archive.

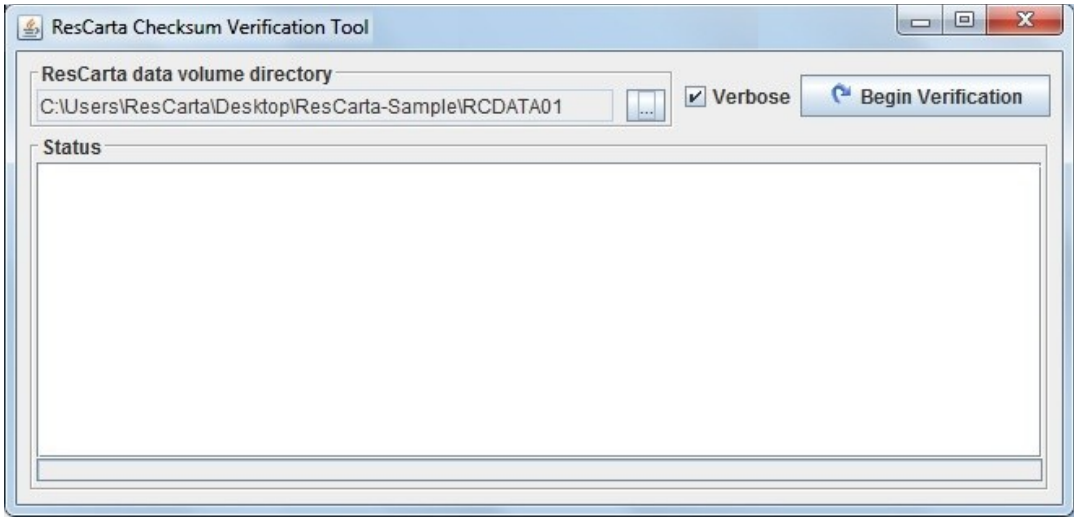
Note: You can use the **ResCarta Checksum Verification Tool** at any time *after the data has been converted into ResCarta format*. For more information on converting image data into the ResCarta format, refer to the **ResCarta Data Conversion Tool** documentation (see page 22).

To verify TIFF checksums of your ResCarta-formatted data:

1. From the **Windows Start** menu, select **Programs**, then select the *Start Menu folder* that you specified during installation (see "To install the ResCarta Toolkit," page 4), and then select **7.- ResCarta Checksum Verification Tool**.

   Note: The version number of the **ResCarta Checksum Verification Tool** is also included as the last part of its **Windows Start** menu option name, but is not included in this document.

   The **ResCarta Checksum Verification Tool** appears.



2. Click the button to the right of the **Source data directory** text field to specify the directory in which your source objects are located. ...

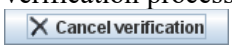
   The Select directory dialog appears.

3. Enter or browse to the **Source data directory** (the directory in which the data for which you want to check accuracy is stored) and click **Select directory**. Select directory

   Tip: You may wish to check both the **RCDATA01** directory on the computer on which your ResCarta archive is stored and the **RCDATA01** directory created by the **ResCarta Data Conversion Tool** (see page 22), to ensure the accuracy of your final and intermediate data, respectively; these verification processes must be run separately.

4. Click **Begin verification**. 

   The progress of the verification process is detailed in the **Status** (lower) section of the **ResCarta Indexer**, and you will be notified of any problems with your data. Checking the **Verbose** option will produce more granular information as the process proceeds.

   Note: The **Begin verification** button is replaced by the **Cancel verification** button during the verification process; click **Cancel verification** at any time to stop the data-accuracy check. 

5. When verification is complete, click the **ResCarta Checksum Verification Tool**'s *close window* button. 

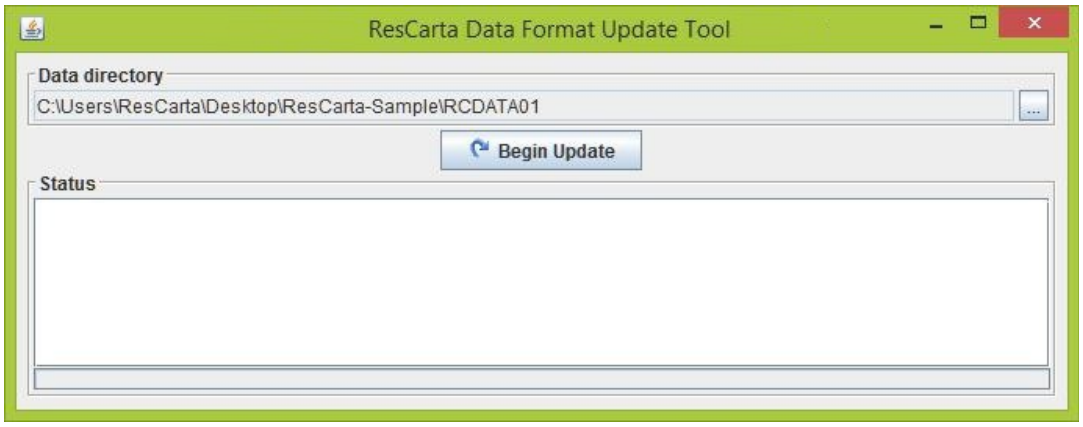
Data Format Update Tool

A utility for bringing older ResCarta Collections up to date with the 6.0 toolkit.

With the release of this version of the ResCarta Toolkit, as with previous major versions, we have updated our METS implementation. In doing this we made changes to the way the ResCarta Collections Manager stores its METS metadata. If you have an existing ResCarta collection with version 1.x or 2.x metadata and want to add to this collection using the new ResCarta Toolkit version 6.0, you will need to run this Update tool ONCE against your existing collection. If you run the tool more than once it will do no harm to your data.

Start the **Data Format Update Tool** and select the location of your ResCarta Data. And press the **Begin Update** button. The status of each file changed will be displayed in the status window frame.

If you have data produced by a 5.0.x version of the ResCarta Toolkit there is no reason to run the Data format update tool. **However once you use the 6.0 Toolkit to modify or create new data you cannot use previous versions of the Toolkit on the archive**.



Toolkit NOTES:

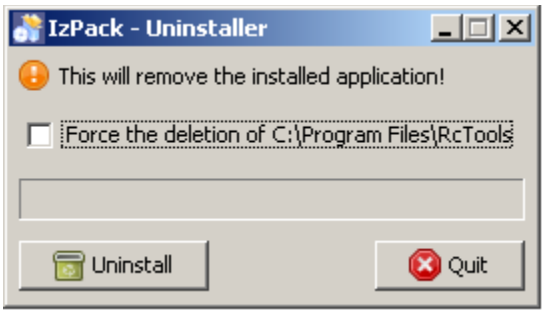
Uninstalling the ResCarta Toolkit

To uninstall any or all of the tools in the ResCarta Toolkit, complete the following uninstallation sequence for the tool or tools you want to remove from your computer.

To uninstall the ResCarta Toolkit:

1. From the **Windows Start** menu, select **Programs**, then select **ResCarta Tools**.

14. From the list select **Uninstall ResCarta Tools**. The tool's Uninstall window opens and the process of removing the tool from your computer begins.



15. If you intend to use ResCarta Tools in the future, please uncheck the "Force the deletion of " the ResCarta tools directory. By doing this you can retain your institutional identifier and the next object id numbers. Click **Uninstall** to continue and remove the software or click Quit to stop and exit the uninstaller without removing software from your hard disk.

A note about Errors:

We work hard to create world-class software, but in the event that you encounter an unexpected error, a dialog box should appear and give you further instructions. If you get a blank dialog box (just an empty dialog) please note what you were doing and send the contents of the error log to the ResCarta Foundation at info@ rescarta.org

This will help us make better tools for you and everyone else. You can find the errors logs under your user directory like…

C:\Documents and Settings\{user name}\.RcTools\logs

Or

C:\Users\{user name}\.RcTools\logs

Glossary

The following terms and definition are used in relation to the ResCarta Toolkit:

Aggregator – The *aggregator* is a required identifier used to limit the number of ResCarta object folders in a single directory within the same source institution parent directory; it must be eight characters (letters and/or numbers) in length. You are required to supply an aggregator (or accept the preselected default) when you use the **ResCarta Metadata Creation Tool** to add metadata to a document.

ALTO - The Analyzed Layout and Text Object (ALTO) XML Schema was initially developed by the METAe project group External URL: http://meta-e.aib.uni-linz.ac.at/ for use with the Library of Congress' Metadata Encoding and Transmission Schema (METS).

Broadcast Wave Format - (BWF) An extension of the WAV audio format and is the recording format of most file-based non-linear digital recorders used for motion picture, radio and television production.

Checksum – A *checksum* is an inspection of numeric values that accompany data, and is used to help determine the integrity of that data.

Destination Directory – The *destination directory* is the directory into which your converted documents will be placed when you use the **ResCarta Data Conversion Tool**.

Metadata – *Metadata* are structured data that describe a document in your collection. For example, title, volume, author, subject, etc.

Monograph – A *monograph* is a publication that's complete as first issued, in one entity. Examples of monographs include books and maps.

Object Identifier – The *object identifier* is made up of the source institution identifier, aggregator, and root identifier. This directory structure (***institution identifier/aggregator/root identifier***) uniquely identifies each ResCarta object in your archive, and is referenced in the **ResCarta Metadata Creation Tool**, the **ResCarta Data Creation Tool**, the **ResCarta Textual Metadata Editor**, and the **ResCarta Collections Manager**.

Parent Directory – A *parent directory* is the directory that is one level up from the current directory.

Raw Transcriptions – An unedited text transcript created by automatically software converting speech into text.

Required Identifier – A *required identifier* is information that must be entered for one or more tools in the ResCarta Toolkit to function.

ResCarta Archive – The *ResCarta archive* is the organized collection of all of your digital objects stored in a known directory structure with open file formats of TIFF and Broadcast Wave along with corresponding metadata in METS (Metadata, encoding and transmission standard) xml format.

ResCarta Object – A *ResCarta object* is any document, book, photo, recording, etc that has been created or converted to a digital representation according to the ResCarta standards. This term is referenced by every tool in the ResCarta Toolkit.

ResCarta Volume – A *ResCarta volume* is a collection of ResCarta objects, stored according to the ResCarta standards. This term is referenced in the **ResCarta Collections Manager**.

Right Click – clicking the secondary mouse button (usually the right button) on a window object.

Root Identifier – The *root identifier* is a required identifier used to distinguish the directory in which a digital object resides; it must be eight characters (letters and/or numbers) in length. You are required to supply a root identifier (or accept the preselected default) when you use the **ResCarta Metadata Creation Tool** to add metadata to a document.

Sequential Naming Convention – A *sequential naming convention* is one in which names are numeric, and increase sequentially. For example, **00000001**, **00000002**, **00000003**, **0000004**, etc.

Serial – A *serial* is a publication that's issued in successive parts, usually numbered, and continuing indefinitely. Examples of serials include periodicals, newspapers, and annual publications.

Serial Monograph – A *serial monograph* is a publication that's complete, or intended to be complete, in a finite number of individual parts. Examples of serial monographs include a three-part book, or a ten-report serial.

Source Data Directory – The *source data directory* is the directory in which the documents you want to convert are stored when you use the **ResCarta Data Conversion Tool**.

Subdirectory – A *subdirectory* is a directory that is one level down from the current directory.

Source Institution Identifier – The *source institution identifier* is a required identifier used to define the institution from which a document originated; it must be eight characters (letters and/or numbers) in length. You can use multiple source institution identifiers with your ResCarta archive. You are required to supply a source institution identifier when you use the **ResCarta Metadata Creation Tool** to add metadata to a document. The ResCarta Foundation suggests the use of an existing registered ID (e.g. Library of Congress MARC organization code, Staatsbibliothek zu Berlin, The British Library,etc.) as the identifier, if one exists for your institution. If the normalized length of the identifier is less than eight characters, then padding with numeric zeros is suggested. You can search the Library of Congress MARC code list at the following url. http://www.loc.gov/marc/organizations/org-search.php You can also add an identifier to the code list by following instructions found at the following url. http://www.loc.gov/marc/organizations/orgshome.html#requests

Transcription software – Software to automatically convert speech into a text transcript

Waveform Audio File Format - (WAV) A Microsoft/ IBM audio file format standard for storing an audio bitstream on PCs. It is an application of the Resource Interchange File Format (RIFF) bitstream format method for storing data in "chunks". The ResCarta Toolkit currently supports mono or stereo 24-bit 96Khz and 16-bit 44.1Khz formats.

Appendix:

With this version of the toolkit we have moved the configuration files out of the install directory to the user's home directory. Recent changes in the way Windows operating systems control the access to the "Program Files" directory have caused problems for some installations. The XML files in the config directory maintain last used settings and input/output directory locations for each of the tools.

The configuration locations will vary with versions of the Microsoft Windows operating systems. They are placed into the user home directory under a "dotted" directory called .RcTools.

Examples

Windows 10 = C:\Users\{user name}\.RcTools\config\

Windows 7 = C:\Users\{user name}\.RcTools\config\

Windows Vista = C:\Users\{user name}\.RcTools\config\

Windows XP = C:\Documents and Settings\{user name}\.RcTools\config\

The RcSystem.properties file

This is a new "advanced" configuration file for the ResCarta Toolkit. If one does not exist in the user .RcTools/conf directory, it will be created with these default values on start of the Data Conversion tool.

Several parameters have changed their default values. In particular the Data conversion tool will no longer resample data to comply with the ResCarta "Book" type specified resolutions. If you are scanning to the suggested resolutions there will be no change in the expected output. If you have scanned materials at more than or less than the suggested resolution, NO resampling will take place unless the **Resampling to ResCarta standard resolutions** parameter is set to "**true**."

Textural Metadata can be carried within the TIF image file tags or in external **ALTO formatted xml** files This is controlled by the ALTO output "True" or "False"

We also allow for non-archive storage of color images. To make JPEG compressed tiff images instead of uncompressed images adjust the **JPEG compressed TIFF output** to "**true**"

ResCarta Data conversion tool now does intelligent conversion of **PDF files**. If the page contains color material it will be converted to a color page. If the page contains only grayscale information, the converter will decide whether to create a grayscale image or threshold the image to a bitone file depending on the values of the **Tolerant bit depth reduction** parameters. Do not adjust these parameters unless you know what you are doing. For more on these parameters, request the "Adjusting bit depth reduction parameters." pamphlet from the ResCarta Foundation.

The **User defined per object pre-conversion JavaScript** sets the location of a JavaScript file that can be used for modifying the output of the ResCarta Data conversion tool. A sample JavaScript file is provided on our website download page. For more on using JavaScript to automate conversions see.

The Mozilla tutorial at http://www.mozilla.org/rhino/doc.html or the Oracle guide at

http://download.oracle.com/javase/6/docs/technotes/guides/scripting/programmer_guide/index.html

```
#-----------------------------------------------------------------------------
# ResCarta System Properties
#-----------------------------------------------------------------------------

# Note: Boolean property settings are true or false (not case sensitive).
# Paths MUST be specified using forward slash characters and NOT backward slash
# characters.

# JAI
com.sun.media.jai.tileCacheSizeMb=256
com.sun.media.jai.disableMediaLib=true

# JPEG compressed TIFF output
org.rescarta.image.tiff.jpeg=false

# JPEG quality setting is a number greater than 0 and less than or equal to 1.
org.rescarta.image.tiff.jpeg.quality=1.0
org.rescarta.image.tiff.jpeg.ycbcr=false

# ALTO    {NEW for  Version 6.0}
#
# If set to false text will stored in TIFF headers.
org.rescarta.alto.output.enabled=true

#-----------------------------------------------------------------------------
# Data Conversion related properties
#-----------------------------------------------------------------------------

# Maximum number of threads to spawn during data conversion
#
# If set to 0, a thread will be spawned for each available core. If set to 1,
# multithreading will be disabled. The data conversion operation will never
# spawn more threads than the number of available cores.
org.rescarta.dct.maxThreads=0

# User defined per object pre-conversion JavaScript
#
# To enable, this property must contain a path to a JavaScript file that
# defines a function named rcPreConvert. The rcPreConvert function must take
# two parameters. The first parameter, of type java.io.File, is the source
# directory or file being converted. The second parameter, of type
# org.rescarta.metadata.RcObjectMetadata, is the metadata for the object being
# converted.
org.rescarta.dct.preConvertJavaScriptFile=

# Resampling to ResCarta standard resolutions
org.rescarta.image.resample=false
```

```
# JPedal PDF text extraction space detection threshold (default: 0.667)
#
# 1.0 equals the width of one space.
org.jpedal.extraction.text.spaceDetectionThreshold=0.667

# Ghostscript output logging
org.rescarta.gs.outputLogging=false

# Tolerant bit depth reduction
org.rescarta.image.bitDepthReduction=false
org.rescarta.image.bitDepthReductionReport=true

# Tolerant bit depth reduction - color tolerances
org.rescarta.image.bitDepthReduction.sampleDeviationTolerance=60
org.rescarta.image.bitDepthReduction.sampleDeviationCountTolerance=10

# Tolerant bit depth reduction - grayscale tolerances
org.rescarta.image.bitDepthReduction.grayHistogramAnalysisXAxisRangeBottom=25
org.rescarta.image.bitDepthReduction.grayHistogramAnalysisXAxisRangeTop=230
org.rescarta.image.bitDepthReduction.grayHistogramAnalysisYAxisRangeBottom=0.003
org.rescarta.image.bitDepthReduction.grayHistogramAnalysisTolerance=0.01

# Tolerant bit depth reduction - write histograms for images with a bit depth
# reduction recommendation of grayscale or binary.
org.rescarta.image.bitDepthReduction.writeGrayHistograms=false
```

Sample JavaScript file

The following JavaScript file will change the aggregators of an existing ResCarta data objects from 00000001 to SK000001. It will add an additional statement "modified by JavaScript" to the existing title of each object and add a note field listing the original path of the object. And finally it will add one additional subject to each object with the authority of "LCSH"

To use this script you must create a file with its contents and add the path to the file in your RcSystem.properties to the parameter called "org.rescarta.dct.preConvertJavaScriptFile=" for example…

org.rescarta.dct.preConvertJavaScriptFile=C:/myjavasrcipts/metadata/RcPreConvert.js

NOTE: the file location above is correct. You will need to use the "/" instead of the windows "\" separator.

RcPreConvert.js

```
importClass(org.rescarta.metadata.RcMetadataSubject);

function rcPreConvert(sourceDirOrFile, rcObjectMetadata) {

        rcObjectMetadata.setAggregator("SK" + rcObjectMetadata.getAggregator().substring(2,8));

        rcObjectMetadata.setTitle(rcObjectMetadata.getTitle() + " modified by JavaScript");

        rcObjectMetadata.setNote("From " + sourceDirOrFile.getPath());

        rcObjectMetadata.getSubjects().add(new RcMetadataSubject(rcObjectMetadata ,"LCSH", "Subject
from JavaScript"));
}
```

Note: See http://software.rescarta.org/apidocs/rc-sdk and Class RcObjectMetadata for more detailed information on variable names.

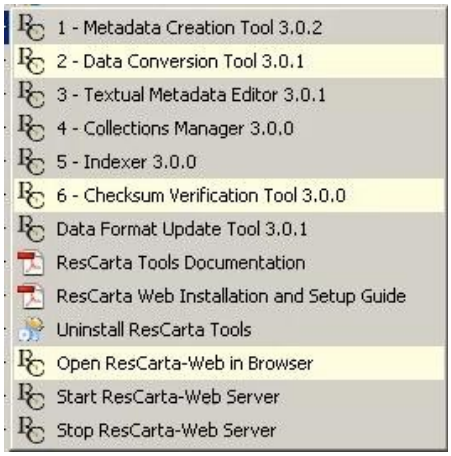
Adjusting tool memory usage

With this version of the ResCarta Toolkit the basic memory has been increased to use one gigabyte of memory. Most modern systems have much more memory available. The memory limit for 32bit Windows XP is 1300 megabytes even if you have more memory in your system. If you need to **reduce** the memory usage of your ResCarta tools due to lack of memory in your system, edit the Rc{toolname}.l4j.ini file found in the default installation directory of C:\Program Files\RcTools-5.0\bin. Change the text from -Xmx1G to –Xmx512M. The G stands for Gigabyte of memory and the M for megabytes of memory. For large audio files with transcriptions it is recommended to INCREASE the memory usage to -Xmx2G or more if your system has > 4G of memory.

ResCarta Tools now support multiple CPU cores and on 64bit operating system can address larger extents of memory.

After installation shortcuts

The new ResCarta Toolkit installer adds a sample data set and the ResCarta-Web application for testing purposes. Under your installation directory you will find an Apache Tomcat server with ResCarta-Web in the Webapps directory. Within the ResCarta-Web directory you will find a RCDATA01 directory with sample data..

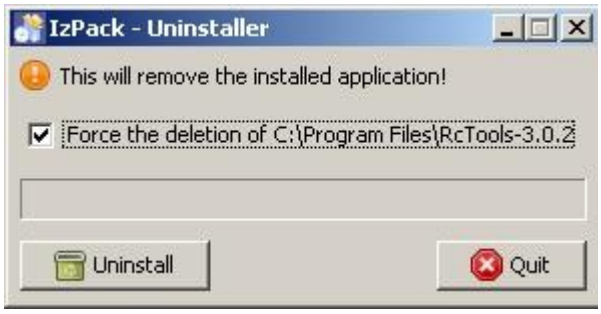


Click the "Start ResCarta-Web Server" shortcut . This will start a TOMCAT window. Do not close the window. This window is the server and it will show messages about the status of the web server,

Once the TOMCAT window is opened you can click on the "Open ResCarta-Web in Browser" shortcut to open your default browser and see the contents of the sample RCDATA01 directory found at the default location of …

C:\Program Files\RcTools-5.0\apache-tomcat-7.0.19\webapps\ResCarta-Web\RCDATA01

You can use this directory for testing the tools and viewing your own collections. But it is advised that you create your RCDATA01 directory in another known accessible location and COPY your data to the sample web application for testing.



WARNING DATA LOSS!

The uninstaller will **delete** all files under the install directory if the "Force the deletion" checkbox is checked. If you have added data to your sample server location UNCHECK this before uninstalling the toolkit.

Using the toolkit from a command line.

The ResCarta Indexer

Java –cp {Classpath} org.rescarta.ir.index.RcIrIndexer{Data path} {Data path}index.ir4 –Parameters

Where

{Classpath} By default would be = "C:\Program Files\RcTools-6.0.X\lib*"

{Data path} Sample= "C:\Program Files\RcTools-6.0.X\apache-tomcat-7.0.67\webapps\ResCarta-Web\RCDATA01"

Parameters can be

-quiet To stop display of indexing performance information as the index is working
-rebuild To completely rebuild the index destroying the existing index

Windows (note the required quotes on directories with spaces)

Java -Xmx1024m -cp "C:\Program Files\RcTools-6.0.X\lib*" **org.rescarta.ir.index.RcIrIndexer** "C:\Program Files\RcTools-6.0.X\apache-tomcat-7.0.67\webapps\ResCarta-Web\RCDATA01" "C:\Program Files\RcTools-6.0.X\apache-tomcat-7.0.67\webapps\ResCarta-Web\RCDATA01\index.ir4" -rebuild -quiet

Linux

java -Xmx1G -cp .:/home/userone/RcTools-6.0.X/lib/* **org.rescarta.ir.index.RcIrIndexer**/data/RCDATA01 /data/RCDATA01/index.ir4 -rebuild -quiet

Generate Thumbnails

Java –cp {Classpath} org.rescarta.tools.cmd.RcWebThumbnailGenerator {Data path} {thumbs path}

Where

{Classpath} By default would be = "C:\Program Files\RcTools-6.0.X\lib*"

{Data path} Sample= "C:\Program Files\RcTools-6.0.X\apache-tomcat-7.0.67\webapps\ResCarta-Web\RCDATA01"

{thumbs path} Sample== "C:\Program Files\RcTools-6.0.X\apache-tomcat-7.0.67\webapps\ResCarta-Web\RCDATA01\thumbs"

Windows

For example if your ResCarta Toolkit is installed in the default location under \Program Files and your ResCarta archive directory is in the root of your C: drive; you would issue the following command on a single line in a command window.

C:\> java -cp ";C:\Program Files\RcTools-6.0.X\lib*" org.rescarta.tools.cmd.RcWebThumbnailGenerator "C:\RCDATA01" "C:\RCDATA01\thumbs"

Linux/Mac

java -Xmx1G -cp .:/home/userone/RcTools-6.0.X/lib/* org.rescarta.tools.cmd.RcWebThumbnailGenerator /data/RCDATA01 /data/RCDATA01/thumbs

Extract textural metadata from Objects

Java –cp {Classpath} org.rescarta.cli.RcDumpText {Data path} {Outputpath}

Where

{Classpath} By default would be = "C:\Program Files\RcTools-6.0.X\lib*"

{Data path} Sample= "C:\Program Files\RcTools-6.0.X\apache-tomcat-7.0.67\webapps\ResCarta-Web\RCDATA01"

{Outputpath} If ommitted will use the Data path else will build a duplicate tree under the output path.

Parameters can be

-n To add newlines after each word

Creates a *.txt output file from each digital object that contains text data. The files are dropped into the original tree or can have a duplicate tree created with text files in the corresponding tree directory.

Windows

For example if your ResCarta Toolkit is installed in the default location under \Program Files and your ResCarta archive directory is in the root of your C: drive; you would issue the following command on a single line in a command window.

C:\> java -cp ";C:\Program Files\RcTools-6.0.X\lib*" org.rescarta.cli.RcDumpText "C:\RCDATA01" “C:\TXT,out”

Linux/Mac dumping from demo data location

java -cp :/Applications/RcTools-6.0.X/lib/* org.rescarta.cli.RcDumpText /Applications/RcTools-6.0.X/apache-tomcat-7.0.67/webapps/ResCarta-Web/RCDATA01 /tmp/TXT.Out

Adding Addtional OCR Language Support

By default the Data Conversion Tool support English as it's OCR language. To support other languages you will need to download a language pack from the Tesseract site. The packs are available in compressed formats. You will need to uncompress them and add them to the ResCarta Toolkit installation directory tessdata directory which by default on windows is C:\Program Files\RcTools-6.0.X\tessdata. To allow for detection of a language, the metadata for the source files must be entered in the Metadata Creation Tool using the preferred language.

Language support is controlled by the contents of your RcTessLangMap.properties file found in your C:\Users\{UserName}\.RcTools\config on latest versions of Windows. You can adjust the file with a text editor. The following shows Russian, French and English settings for the RcTessLangMap.properties file. Note that French and English ISO language codes do not map directly to the Tesseract codes.

```
#Map ISO languages to Tesseract languages for ResCarta data conversion
#Mon Nov 23 12:23:17 CST 2015
rus=rus
fre=fra
eng=en
```

You can find language packs for Tesseract at the following url...

https://github.com/tesseract-ocr

NOTES: