UNIVERSITY LIBRARIES
UNIVERSITY AT ALBANY
State University of New York

IMC
Innovate Make Create
http://library.albany.edu/imc/
518 442-3607

Scanning Text to *<u>EDIT</u>* a Document
Use OmniPage Pro on Workstations 2, 18, 25, 27 and 32
# includes converting a PDF file to a Word file

The copyright law of the United States (Title 17, United States Code) governs the reproduction of copyrighted material. The person using this equipment and software is liable for any infringement.

OmniPage Pro will <u>NOT</u> make an exact replica of your document.

From the Start menu select **OmniPage Professional** to create an "edit-able" document. If you have a text document and want to preserve its look exactly or if you want to use it on the Web, use Adobe Acrobat to create a PDF file.

Optical Character Recognition (OCR) transfers paper documents into edit-able computer files. Your original document is scanned or imported as an image. OmniPage Pro interprets it as text or graphics. In general, all text is interpreted as ASCII characters (text); all other marks (hand written notes, smudges or pictures) are zoned as graphics. OmniPage Pro recognizes foreign languages, tables, multiple columns, etc. Output formatting can be manipulated Following are basic instructions.

1. Use the ADF, Automatic Document Feeder or lift the cover on the scanner and align the page to be scanned. Close the cover. **NOTE:** The program defaults to the settings the previous user selected. You may override the settings by selecting an option listed in the dropdown menus by each step. Additional settings are available under **Tools > Options**

2. Select the following windows to view your work. From top line menu select **Windows > Classic View**.
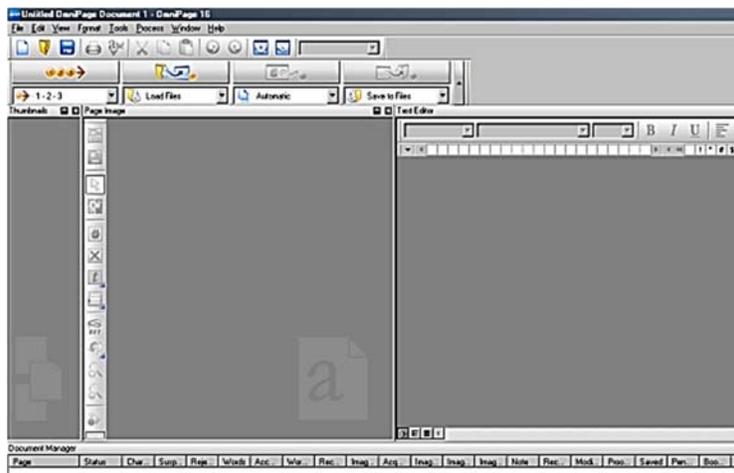
BEFORE scanning or loading an image file, consider the following:

- The program defaults to the settings the previous user selected. You may override the settings by selecting options listed in the dropdown menus. Additional settings are available under **Tools** > **Options.**
- If using the Scanner, select **Tools** > **Options > Scanner.** Select the source of your document.
- Do not retain graphics when saving and exporting (see converter option). Import and place graphics separately as a last step. It is easier to import an image and manipulate its placement later than to include it with the text now. Scan images from another application.
- If you are importing multiple columns, export into one column. Reset columns after editing.
- Cleanup within the word processing application you are saving the file to. PROOFREAD! Be prepared to edit all text. This is not perfect. For example, OCR may assume the number one (1) is a lower case L.
- If you chose to proofread finished text or if it was already selected, the **OCR Proofreader** window will open after you scan and the OCR Proofreader selects all unknown/unsure words for decision making. Highlighted is what you see; the suspect word/character is OmniPage's interpretation. Use like spell-check; see Tools > Options > Proofing
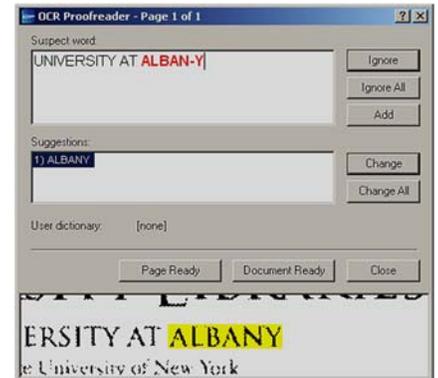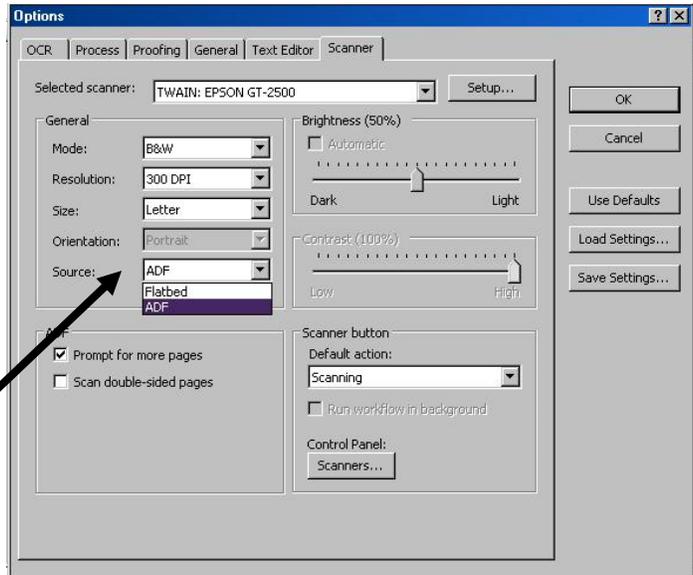
Begin by reviewing the selections on the Toolbar. **1-2-3 OCR.** Settings are commonly set to:

1. Document Source: Scan Black and White
2. Original Layout: Automatic
3. Export Results: Save to file

To convert an image file such as Acrobat PDF to a file for editing purposes using applications such as Microsoft Word, Excel, etc. From the top-line menu, Select **Load Files** from the second dropdown of step 1 of 1-2-3.
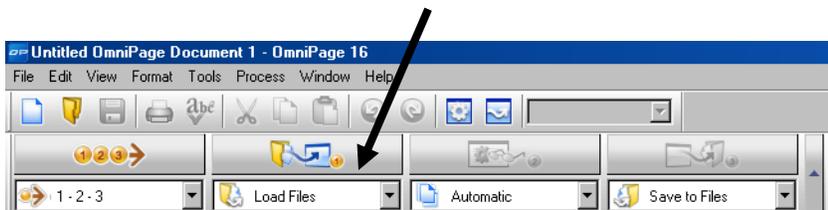
Proceed to scan by selecting:

3. The **Save** to file box automatically displays after OCR is applied and your document appears in the Text View or after the OCR Proofreader is finished. You may select Cancel, continue scanning, and save later.

   OPTION: If you are interrupting your scanning project and want to continue at a later time - or – want to scan and then perform OCR Proofreader on <u>all pages</u> at once, save the document as an OmniPage Pro document. From the top line menu: **File > Save OPD**. Later, when you return, select **File > Open OPD**. Check settings and continue. When finished, save the file. When the file is saved as an OPD file, it must be exported as a different file type. Select **File > Export Results > Save to file.**
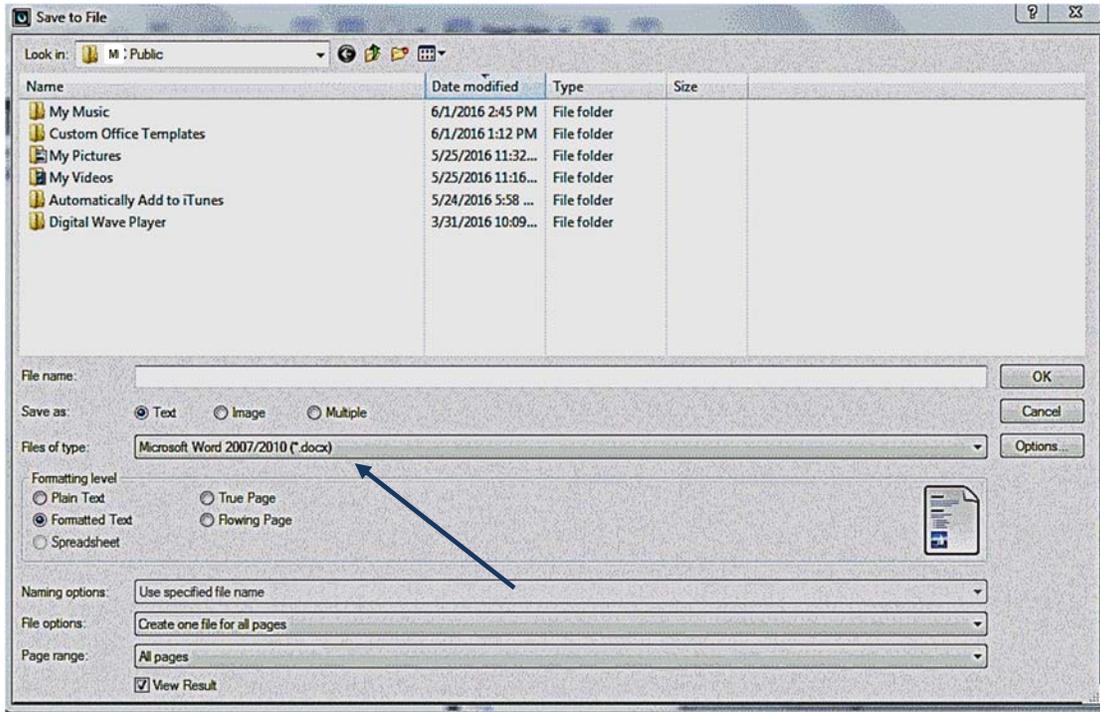
4. **NOTE! Export/Save to file formats. Select the output format of your document.** Select the settings that specify how much formatting of the original document you want to retain. When in doubt select Save to File Choose a file type.

   Files of Type Select the correct version of the application you are using; select Microsoft Word 2007/2010 for Word versions after 2010. Select save and Launch; the file displays.

TIP: Save the document with different Formatting Levels using different file names. Compare later what file is best suited for editing. Select Step 3, Save to Files, only after scanning to repeat for multiple files.
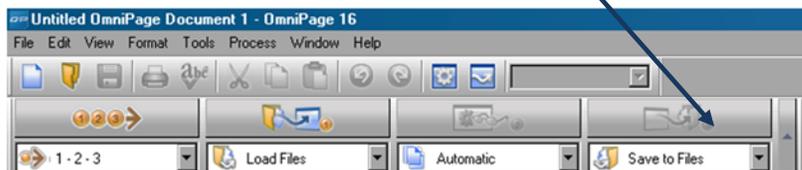
Formatting Level: If you selected **Save as: Text**, select from the following**:**

Plain Text: select if decision is to strip the document of all formatting. All recognized text is one font type, one column and left aligned. Suitable for ASCII, plain text and formats with extension .txt or .rtf Use this formatting level if you are capturing only the text and have much editing to do.

Formatted Text: exports one column of text, retaining font characteristics and some paragraph formatting; will retain. **This is the most typical selection.**

True page: Select to retain the original appearance of the scanned document. Select if very little or no editing is to be accomplished. All page elements are placed in frames. This is suitable only for file types and applications capable of handling frames or text boxes. Must select for XML export and for all PDF exports except to the file types "PDF edited."

Flowing Page: Select if very little or no editing is to be accomplished. Select this to keep the original layout-especially when using multiple columns and indent settings--not with frames. Text will flow from one column to the next.
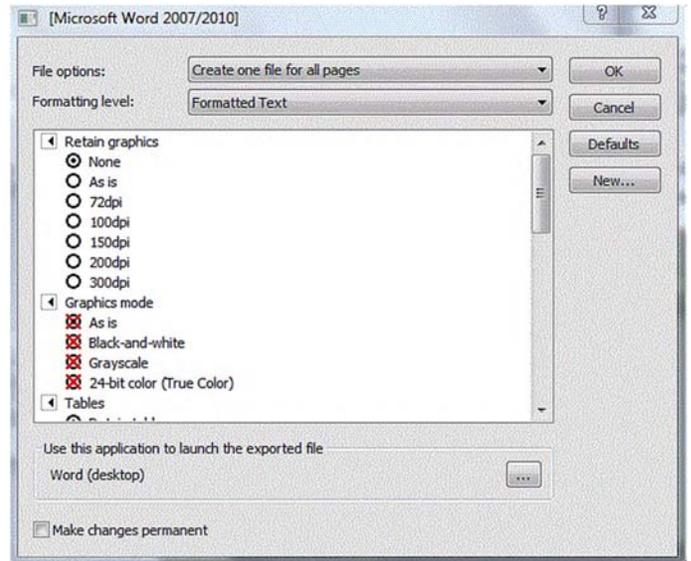
Select **Converter Options**. The formatting level you selected corresponds to the options selected or not selected. Select the formatting level first (see step #4) and then make decisions. Select **apply** if you made a change.

Retain graphics select "None." For best results, scan desired graphics using the Image Scanning application.

5. **File options** set to "**Create one file for all pages**"

6. **Continue with Save. Exit** from Text Scanning when done or press →
    to continue.