

Guide to Creating Accessible Portable Document Files (PDF)

U.S. Department of the Interior

Creating an Accessible Portable Document Format (PDF)

Background:

In 1998, Congress amended the Rehabilitation Act to require Federal agencies to make their electronic and information technology accessible to people with disabilities. Called Section 508, this law applies to all Federal agencies when they develop, procure, maintain, or use electronic and information technology.

Inaccessible technology interferes with an individual's ability to obtain and use information quickly and easily. Section 508 was enacted to eliminate barriers in information technology, to make available new opportunities for people with disabilities, and to encourage development of technologies that will help achieve these goals.

The PDF was developed by Adobe, Inc. in the early 1990's to allow electronic documents to be shared among computer users without the need to have the originating software installed. Its' easily shared format and the availability of the reader software has made the PDF a very popular means to distribute information.

Putting 508 into Action:

Many people require the use of screen reader programs such as Jaws or Window-Eyes to read electronic files, including PDF's, on their computers. These programs literally read aloud what is on the computer screen.

Federal Web managers have been mandated to make their Web sites 508 accessible to individuals that use screen readers so they can obtain the information they need. Historically, PDF's were left out of this requirement but recent Department of the Interior policy has mandated that ALL PDF documents intended for distribution via Web *or e-mail* be created so they can be read using the screen reader technology.

This guide offers an introduction to using MS Word in conjunction with Adobe Acrobat software to create accessible PDF's. Additionally you will be able to:

- Explain why creating accessible PDF's is important;
- Cite examples of MS Word features that will create better PDF documents;
- Create a pdf file from a Word document;
- Add, delete and modify pages in a pdf file;
- Add properties to a pdf file; and
- Modify a pdf file so that it can be read properly by a screen reader.

Most Adobe PDF documents begin on computers as Microsoft Word documents; then are converted to PDFs at the end of the process for distribution. This guide explores how to produce a Microsoft Word document so that, when exported to an Adobe PDF document, it will allow screen readers to correctly and completely read the document.

There are many elements that affect accessibility in a document, such as color use and contrast. However, the main focus of this guide is to make you aware of how to format Microsoft Word documents so that screen readers can read your exported PDF files properly. For more detailed information about colors and contrast, please reference the Other References page at the end of this document.

Instructions for creating accessible PDF's from scanned documents are also included in this guide.

Tips for Structuring a Microsoft Word Document for an Accessible PDF Export:

The following are suggestions for structuring a Word document for easier conversion to PDF.

- o Keep Your Word Document Layout Simple: Keep all of your content in a linear single column progression. This will allow screen readers to obtain a proper reading order of your document. More complex document layout design can scatter the document's reading order for screen readers. Complex documents will need to be corrected manually after conversion by using the TouchUp Reading order feature of Adobe.
- o Use Styles Instead of Text Attributes: To provide structure to your document use the style formatting tool instead of the text bolding feature for text editing. Some screen readers can read out the document's heading labels along with the document's text. For example, the reader will read out loud "Heading 3" for the heading above this paragraph. These headers will become bookmarks in the PDF.
- o Create Alternative Text for Photographs or Graphic Images within Microsoft Word: Right mouse click your document image and select the Format Object from the popup menu and click the Web tab to type an alternative text passage in the text field describing the image.
- o Add Hyperlinks to your Documents: If you want the screen reader user to access a web link from your PDF, use the Microsoft Word's Hyperlink tool by selecting the **Insert** menu and choose the **Hyperlink...** menu to hyperlink a web address.
- o Use Standard Text Fonts: Use standard text fonts such as Times New Roman and Arial.
- o Data Tables: Avoid nested tables. Use the Insert – Table option instead of creating a table with text boxes or layers.
- o Read the Table guidelines tips section found in the Other References section at the end of this document.
- o Add a Blank Page to the End of Your Document: Make sure you add one blank page at the end of the word document or else the accessible screen reader may

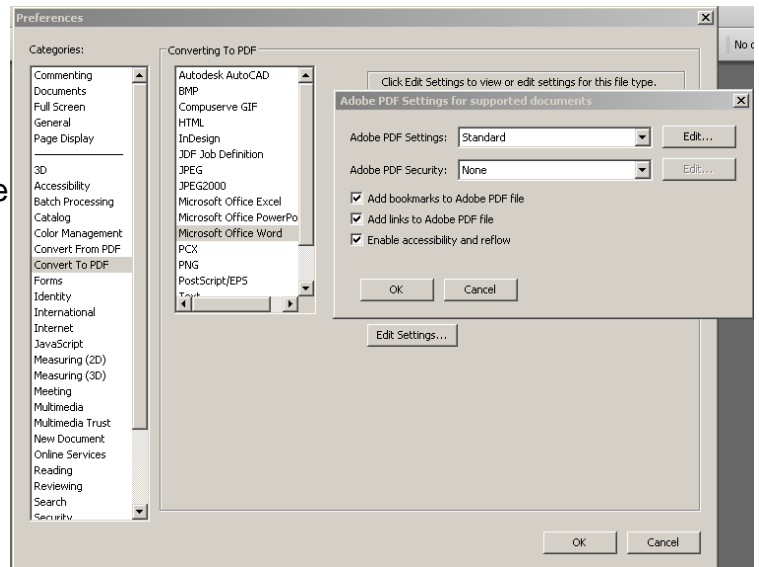
- o not read the last part of your document.
- o Once your document has been converted to PDF, check the tabbing order to be sure that the document reads as intended.

Things to Note:

- o When using Roman numerals, screen readers will read Roman numerals as a letter instead of a number equivalent. Instead of using Roman numerals, use numbers.
- o Use the Bullet and Numbering feature from the Format menu when bullets are required. Some screen readers will read out the word "bullet" to the user if the document is formatted properly.
- o Use proper punctuation to allow for natural pauses and breaks in your lists. Screen readers will pause when coming to a punctuation mark.
- o Information in the Header and Footer are NOT read by screen readers once the document is converted to PDF.
- o Avoid using Word based check boxes and blank lines for forms since the screen reader cannot detect these items for the end user. Once the document is converted to PDF, use the forms feature in Adobe Acrobat to create the form fields.

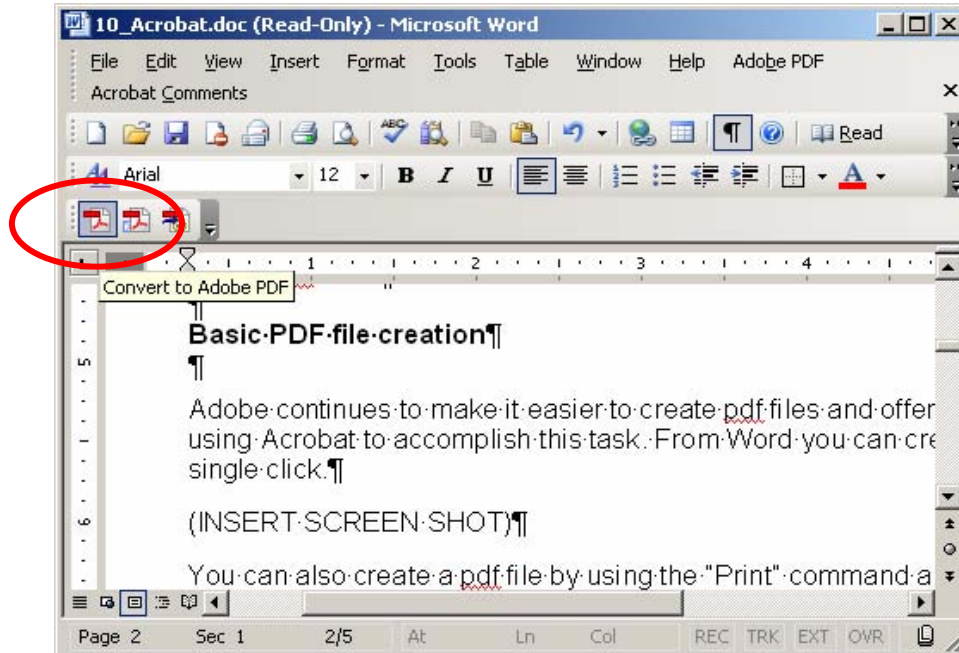
Adjust Adobe Acrobat's Preference Settings: Before converting a Word document into a PDF file you must adjust Acrobat Professional's Preference settings with the following steps. (The following is shown for Adobe Acrobat 8 and may differ depending on the Adobe version you are using)

1. Open Adobe Acrobat.
2. Choose **Edit and Preference** from the top menu.
3. Choose **Convert To PDF** in the Categories field and select **Microsoft Office Word** in the Converting To PDF field, and select the **Edit Settings...** button.
4. In the 'Adobe PDF Setting for supported documents' menu box choose **Standard** from the Adobe PDF Settings drop down menu. (Note: see trouble shooting section if accessibility is not available in the drop down menu)
5. Check the Add bookmarks, Add links and Enable accessibility & reflow boxes.
6. Finally, select the **OK** button to save each one of your menu preference box changes.

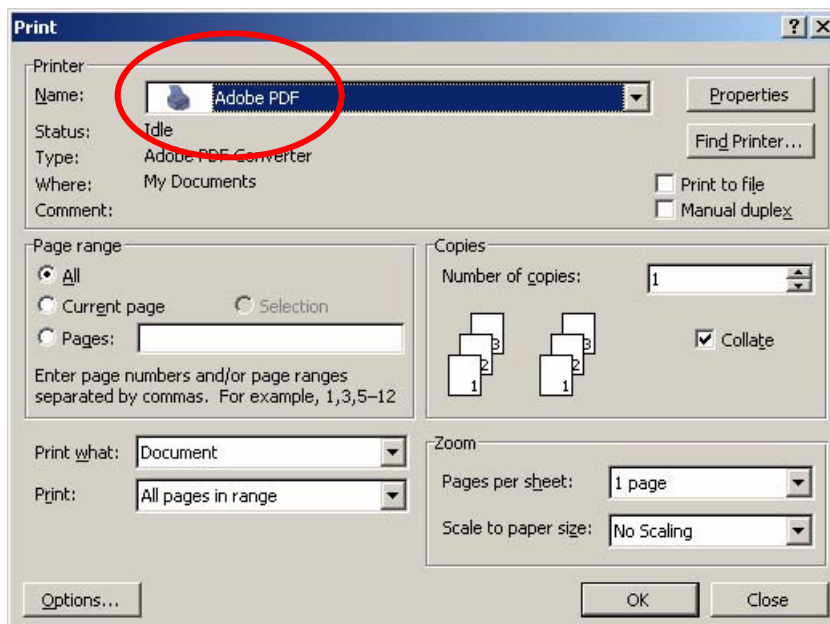


Basic PDF file creation

Adobe continues to make it easier to create PDF files and offers multiple ways of using Acrobat to accomplish this task. When you have Adobe Acrobat installed properly, you can create a PDF file with a single click directly from Word.

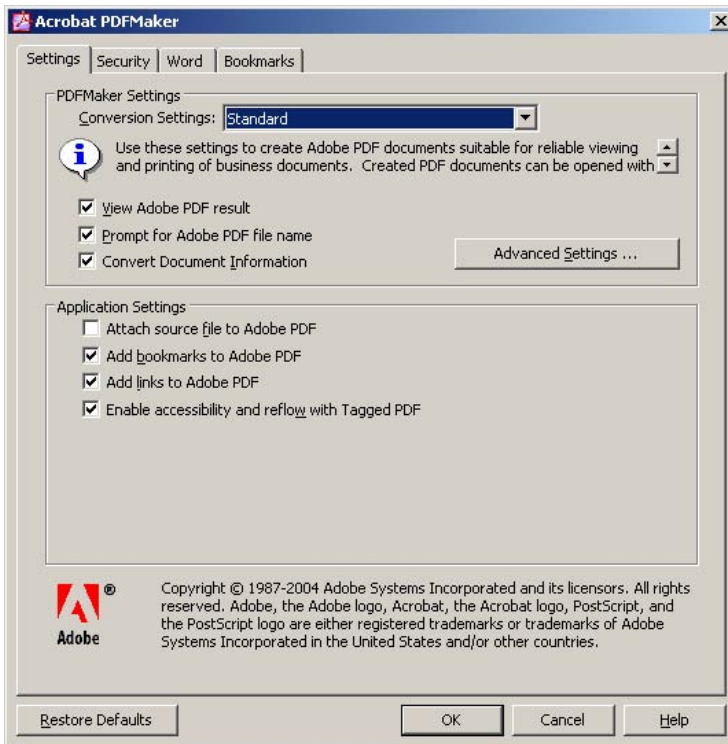


You can also create a PDF file by using the "Print" command and selecting "Adobe PDF" as your printer. However, be aware that some features do not convert from MS Word when creating PDF's this way.



Optimizing PDF Files

The versatility of PDF allows the format to be used on the Web, where small file size is important, and also in the print industry, where the quality of the finished product is much more important than file size.



When creating a PDF, you should take steps to optimize the PDF for the appropriate purpose. Acrobat has some built-in settings that you can use or modify to create your own. In Word, go to: Adobe PDF > Change Conversion Settings

Click “Advanced Settings” to customize the conversion settings.

Modifying the PDF File

1. Add Properties

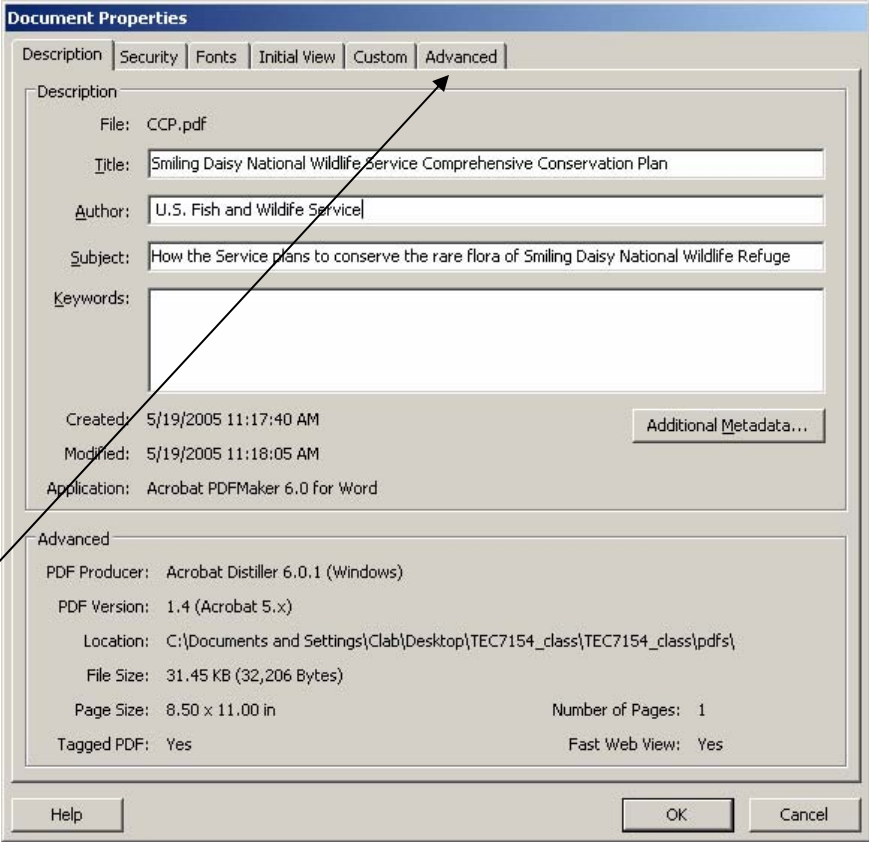
Once your PDF file is created, you’ll want to add “properties” that describe the file. These properties go with the PDF file and may be used by search engines to discover and reference the PDF in searches. The language specification is referenced by accessible technology.

The document properties screen is found under File - > Properties. At a minimum, you should always add the following properties to your document.

- Title
- Author
- Subject
- Keywords
- Language

Note: Some of this information could have been carried over from the original source document, which may not be appropriate.

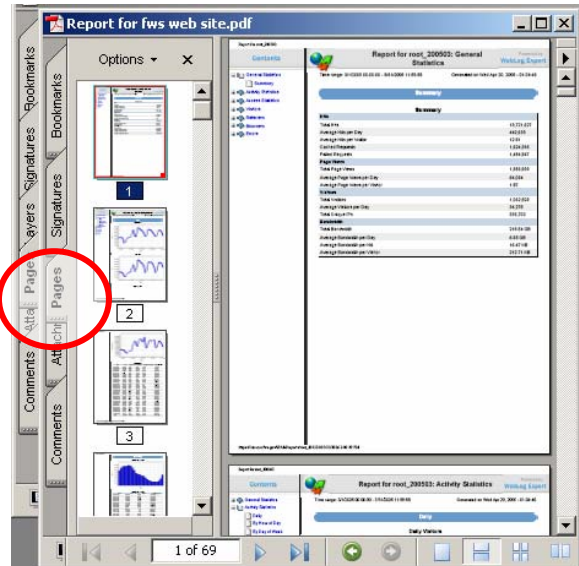
The language specification field is found on the Advanced tab of the Properties screen under reading options.



2. Delete, Rearrange, Insert Pages

Along the left side of the Acrobat window are a series of tabs. See the “Pages” tab? If not, select pages from View > Navigation Tabs > Pages.

Click on the pages tab to see a thumbnail of each page in the PDF document. From this window you can rearrange or delete pages, extract one to create a separate PDF file, or insert another PDF file.



3. Touch Up Text

Acrobat allows limited editing of PDF files. If you need to change a word or two, correct a number, or make similar changes, go to: Tools > Advanced Editing > TouchUp Text Tool. Select the text that you need to change, and type in the new information. If extensive changes to text is required, it is suggested that you access the original Word document to make those changes and then convert the new document to PDF.

4. Changing a Scanned Document to Real Text

A PDF file made from a scanned document, such as a signed memo, may look to contain text, but it is really just a picture of text. Zoom in on a portion of the text. If you see jagged edges or other imperfections, you're probably looking at a scanned document. Try using the TouchUp text tool. If you can't select text, it's probably just a picture of text. In order to be able to edit the text, or in order for a screen reader to read the text, the text must be recognized from the scan and converted to real text.

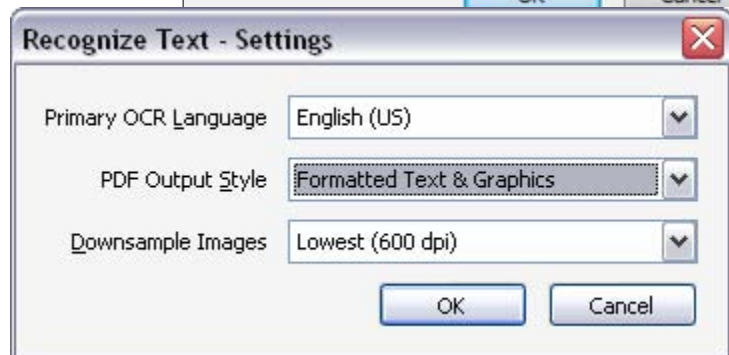
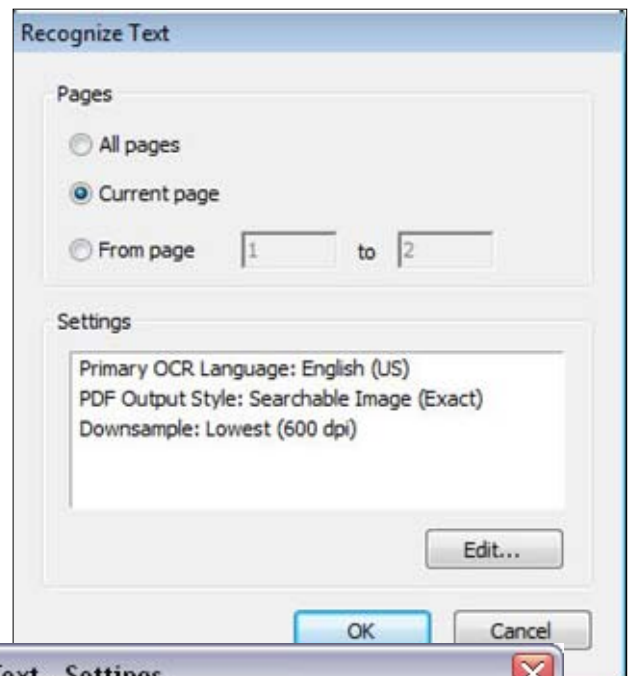
This is what "optical character recognition" (OCR) is. A computer program, in this case Adobe Acrobat, looks at the scan and determines what characters are being depicted. It then replaces the image of the text with editable text. PDFs created from scanned documents are frequently the biggest offenders in regards to accessibility. To correct this:

Go to: **Document > Recognize Text Using OCR > Start**

In the "Recognize Text" dialogue box, you can choose which pages you want recognized. More importantly, you can choose how you want Acrobat to output the text to the PDF file.

To change the output settings, click on the Edit button in the recognize text window. Choose the "Formatted Text & Graphics" in the PDF Output Style to create a screen readable document.

Acrobat can retain the image and hold the text "behind the scenes." This allows the search function to work, but doesn't allow you to edit text or for a screen reader to read the file. Or it can replace the text image with real text. This is preferable for accessibility, editing, and file size; but it can require a lot of cleanup.



Finding Suspects:

After the OCR has converted the image to real text, you must review the document for “suspects”. These are instances where the text image wasn’t clear enough for Acrobat to make a good interpretation. You should progress through the document, suspect by suspect, and correct Acrobat’s errors. To see and correct suspects, go to:

Document > Recognize Text Using OCR > Find First OCR Suspect

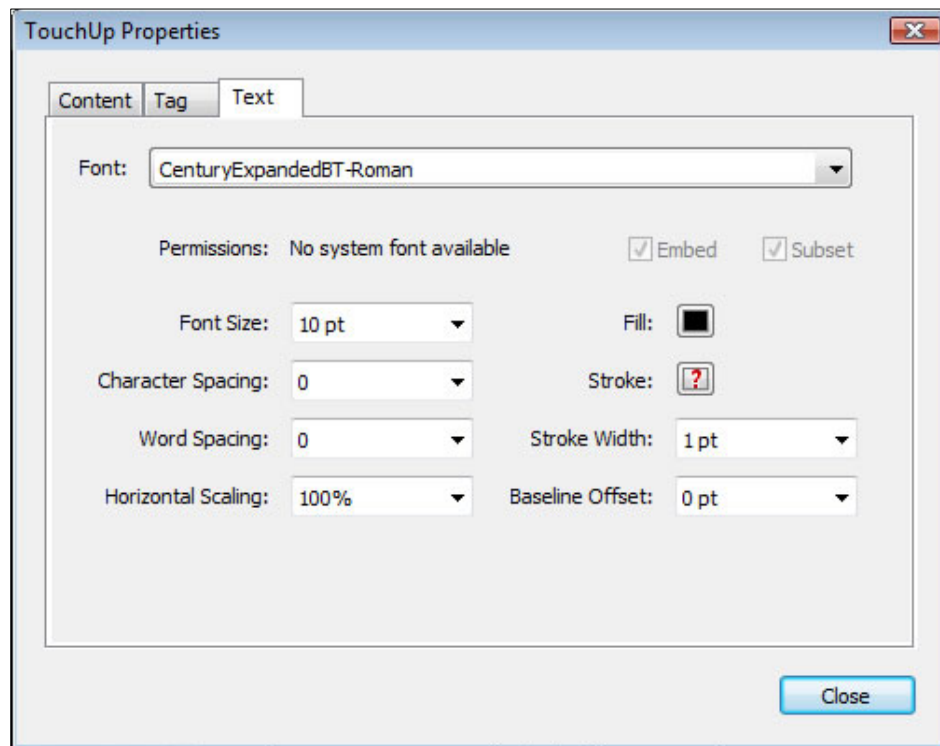
Acrobat will offer a suggested reading. You can accept that or type in your own.

As you’re working through this task, you may find another problem: Acrobat has recognized some text but put it in the wrong font. This happens typically with poor scans or poor originals (photocopies, faxes, etc.).

To fix this, you must change the properties for that bit of text.



First, select the text using the TouchUp text tool. Then, right-click (CTRL-click on the Mac) and select Properties from the resulting drop-down menu. In the properties window, change the font, size, and any other properties to give the selected text the appropriate characteristics.



5. Make a PDF Document Accessible – final steps

There are two things you need to do to make a PDF file accessible to people who need to have the document read aloud.

1. You must establish a structure and reading order so that Acrobat knows how to read through the document.

Acrobat is willing to take a stab at establishing the reading order automatically.

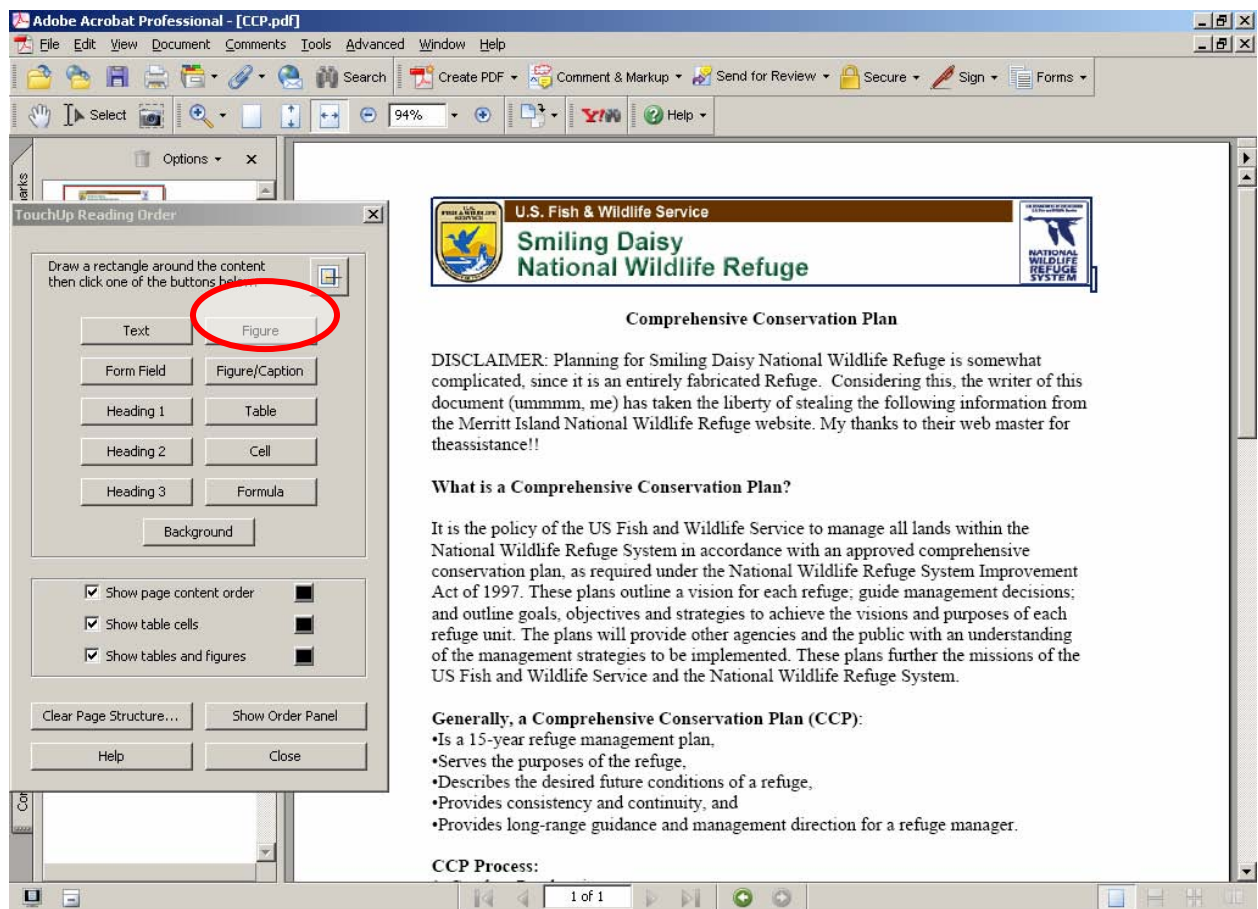
Go to: **Advanced > Accessibility > Add Tags to Document**

Unfortunately, this doesn't work very well if the document has multiple columns, photos, and captions. You'll probably find yourself having to redo much of Acrobat's work manually. Fortunately, while it can be tedious, it's fairly simple work.

Work through the document page-by-page. First, set the reading order.

Go to: **Advanced > Accessibility > TouchUp Reading Order...**

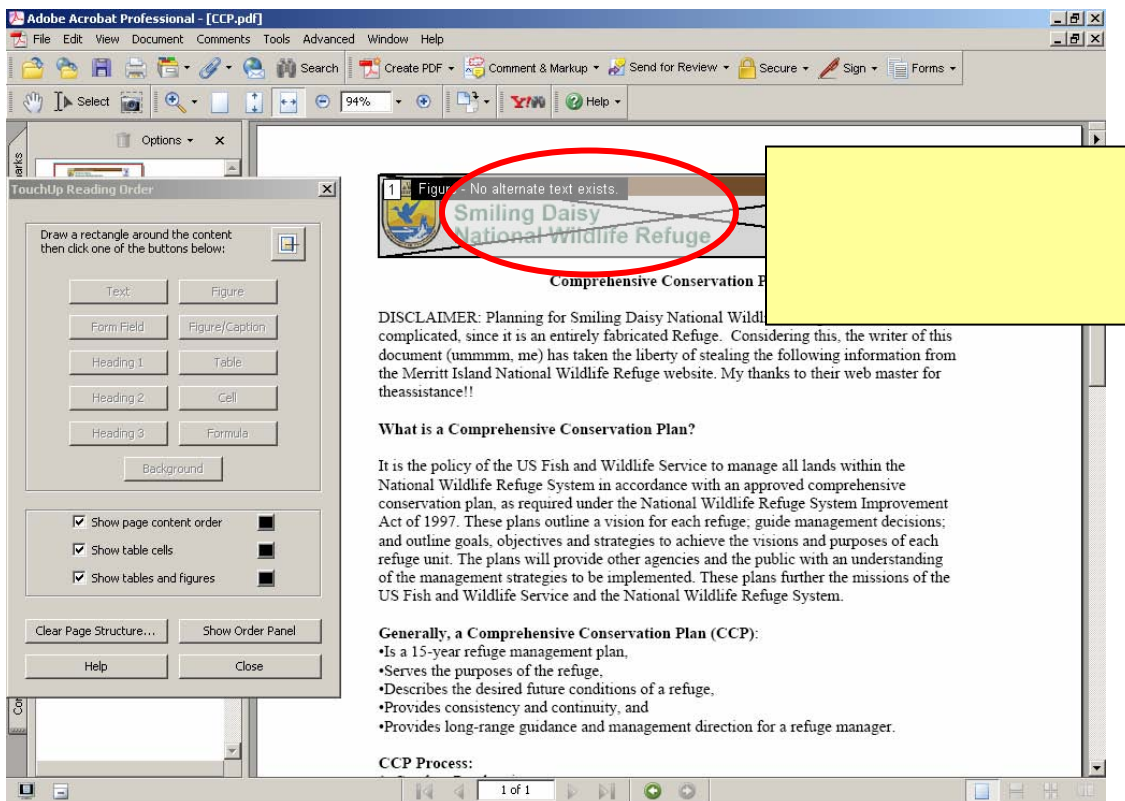
Then use your mouse to draw rectangles around the page content in the order in which the content should be read. From the TouchUp text window, click the button that describes the content .



2. You must give a textual description of non-text items such as graphics and charts. If you added Alternate text to your graphics in Word, those descriptions should have converted to the PDF.

What about graphics? If the graphic is purely decorative, you may want to skip it altogether. If it is not referenced at a specific point in the text, you may want to include it in the reading order at a convenient point, for instance, at the end of a paragraph. For graphics with captions, enclose them both within a box and click the “figure with caption” button.

Using an ALT tag. The ALT tag gives a text description of the figure. Once you’ve placed the figure into the reading order, right-click (ctrl-click on the Mac) inside the box containing the figure. From the drop-down menu that results, select “edit alt tag.” Enter a concise description of what the figure depicts. This takes a bit of artistry, as you must decide what information content there is in the figure that would be of value to someone who can’t see it.



Proceed in this manner until you reach the end of the document. To test your work, have Acrobat read the document back to you.

Go to: **View > Read Out Loud >** and choose how much of the document you want read back to you. For testing purposes, you should probably stick to reading one page at a time. Return to the TouchUp text reading order tool if you need to re-do a page.

Other References:

1. Adobe's Web Site:
http://www.adobe.com/enterprise/accessibility/pdfs/acro7_pg_ue.pdf or
<http://www.adobe.com/enterprise/accessibility/training.html>
2. WebAIM's Web Site: Alternate Text for Images Tutorial
<http://www.webaim.org/techniques/word/>
3. National Center on Disability & Access to Education:
4. Microsoft Word's Tip and Tools for Accessibility (Featuring Tables)
<http://ncdae.org/tools/factsheets/word.cfm>
5. University of Wisconsin: Microsoft Words Structural Markup Tutorial
<http://www.cew.wisc.edu/accessibility/tutorials/structuralmarkuptutorial.htm>
 Table guidelines tips as quoted:
 - "If you have one row of headers across the top, you can set this row as a header by selecting the table and choosing Table >Table Properties > Row tab > Repeat as Header Row at the top of each page."
 - "If a table has more than one row of headers, or it has a set of column headers, it is not possible to add proper headings." (NCDAD Web site, 12/6/2006)."
6. IBM's: Creating Accessible Microsoft Word's Documents. (Featuring Tables)
<http://www-306.ibm.com/able/guidelines/documentation/docmsword.html>.
 Table guidelines tips as quoted:
 - Do not use tabs or spacing to create tables. While it may visually look like a table, it will not be recognized as a table by assistive technology and will not be accessible
 - Simple tables created in Word using Table-Insert-Table are accessible without any modification.
 - Complex tables created in Word using Table-Draw Table may not be accessible because they usually have cells of different heights or a varying number of columns per row. There is no function provided in Word to associate the row and column headers for complex tables
 - Do not use nested tables.
 - Add a caption above or below the table to summarize the purpose of the table.
 - Do not use patterned backgrounds on tables," (IBM's Web site, 12/6/2006).

