PDF White Paper - Final Revised – February, 2010

Updated by 2010 PDF Work Group of the EPA Web Council
Charles Bert, Judy Dew, Ross Gilleland, Michael Hessling, Jaclyn Kelly (BAH), Phoebe Stinson (BAH), Johnny Zimmerman-Ward (E² INC)

Note: Any instructions in this paper refer to Adobe Acrobat 9 Professional unless otherwise stated.

Section 1 - Introduction

PDF is a widely used format for presenting documents electronically that looks exactly like the printed originals. The PDF file format is an ISO standard.

The purpose of this paper is to provide an overview of the many issues EPA staff may encounter when working with PDFs, specifically when making them web-ready for the public. This document discusses each issue and provides solutions either as guidance for best practices or as official policy and/or standards.

The original PDF white paper was developed in 2006 by the US EPA Web Workgroup (WWG) PDF Subcommittee. In 2010, a workgroup from the EPA Web Council reviewed and updated the white paper. While the primary issue was file size limits for posting, the workgroup reviewed and made revisions to the entire white paper.

There are EPA Standards for:
- Linking to a PDF
  http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-linking
- When to convert a document to PDF
  http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2
- Metadata required in a PDF
  http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2
- EPA no longer requires or recommends the division of large PDF documents into smaller ones. If you decide to divide one anyway there are cover page requirements.
  http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2

This workgroup recommends:
- Using Adobe Acrobat Professional because of the accessibility components (but other PDF tools are allowed).(Section 10)
- Using the most recent version Adobe Acrobat Professional because compression and accessibility have improved in each version (Section 10)
- Saving to the most recent reader minus one. (Section 9)
- Use the original clean electronic document for conversion to PDF. Don't scan documents if an electronic version is available. Scanned documents are not inherently accessible. (Section 8)

Best Practices included in this paper
- When to use PDFs (Section 2)
- Making PDF Documents Accessible (Section 3)
- Linking to PDFs (Section 4)
- Making Large Documents User Friendly (Section 5)
- Metadata (Section 6)
- Techniques for Reducing PDF file size (Section 7)
- Dealing with Scanned PDFs (Section 8)
- Adobe Reader Compatibility (Section 9)
- Which Version of Acrobat Professional Should You Use (Section 10)
- Security Features (Section 11)
- How to Create Web-ready PDFs from Word, Excel, and PowerPoint files (Appendix C)

Section 2 - When to use PDFs

See EPA's PDF Standard - When to Use PDF, Document Metadata, Info Required for PDF Sections
http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2

Section 3 - Accessibility

Why are PDFs an accessibility problem

PDFs generate a picture of how a printed document looks; this makes PDF a useful format for circulating picture-perfect replicas of printed documents electronically. This is extremely useful when you need to preserve the look of a document, whether because of graphics and design or because of official pagination and formatting. Unfortunately, because a PDF is *a picture of a page of text* rather than actual text, PDFs are not inherently accessible to visually impaired people or search engines.

You can tag PDFs (using Adobe Acrobat Professional) so that screen readers will be able to recognize the content and read it aloud. This is much like adding alt text to graphics in HTML. Depending on your document, tagging a PDF may go very smoothly. For some documents it is practically impossible (or exorbitantly time-consuming).

The 2010 work group recommends getting the latest version of Adobe Acrobat Professional**.** The group also recommends keeping current on the Adobe Acrobat Professional versions because the file compression and accessibility tools have improved with each release.

Note: Do not use *File > Print > Adobe PDF* that you used in previous versions of Acrobat unless the Adobe PDF menu is unavailable in your authoring software. Making the document accessible is harder and you lose the metadata from the original document.

Note: Appendix C provides directions on how to prepare Microsoft product files for conversion to PDF. It is easier to add information to the Microsoft product files to ensure accessibility than it is to repair a PDF for accessibility.

How to make a PDF accessible (summary)

1. **Tag Your Document**: (NOTE: You can skip this step if you converted your original document as a tagged PDF, e.g. in Word.)
   o **If you are creating the PDF from a Microsoft Office application,** choose *Adobe PDF > Change Conversion Settings* and select "Enable accessibility and reflow with Tagged PDF". Then select OK. Now create the PDF using *Adobe PDF > Convert to PDF* (or the Convert to PDF icon).
   o **If you are using an existing PDF**, tag the file for accessibility using Adobe Acrobat Professional (*Advanced > Accessibility > Add Tags to Document*)

2. **Run a Full Check**: (*Advanced > Accessibility > Full Check*), then look at the report it generates and see what you need to fix.
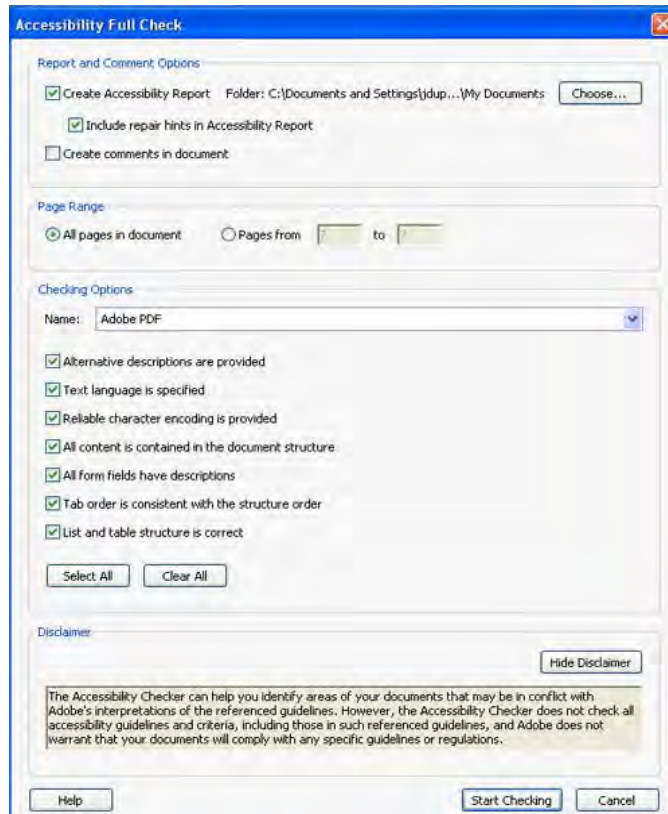
3. **Check the Reading Order**: Check the reading order and correct what doesn't make sense (***Advanced > Accessibility > Touch Up Reading Order***)
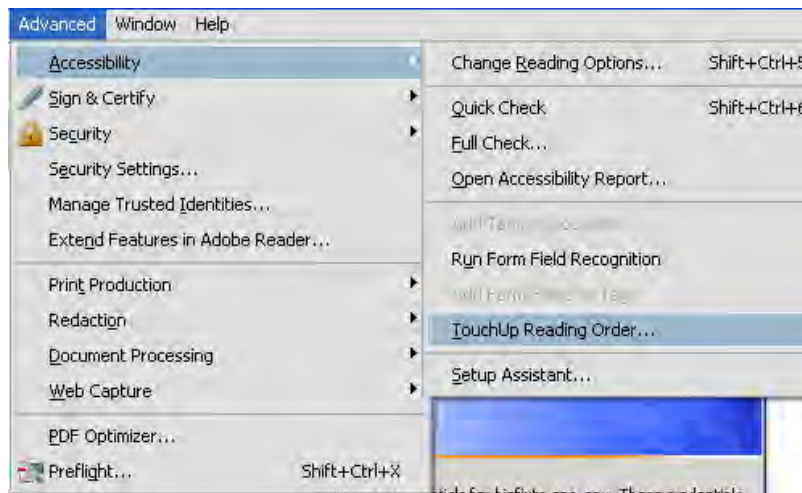
4. Add alternate text for images (***Tools > Advanced Editing > TouchUp Object Tool***). Select the image that needs alt text. In the Tags Options drop down menu select ***Find Tag from Selection***. Right click on the image tag and go to properties. Enter the alternate text in the Alternate Text field. This can also be used to edit alt text.
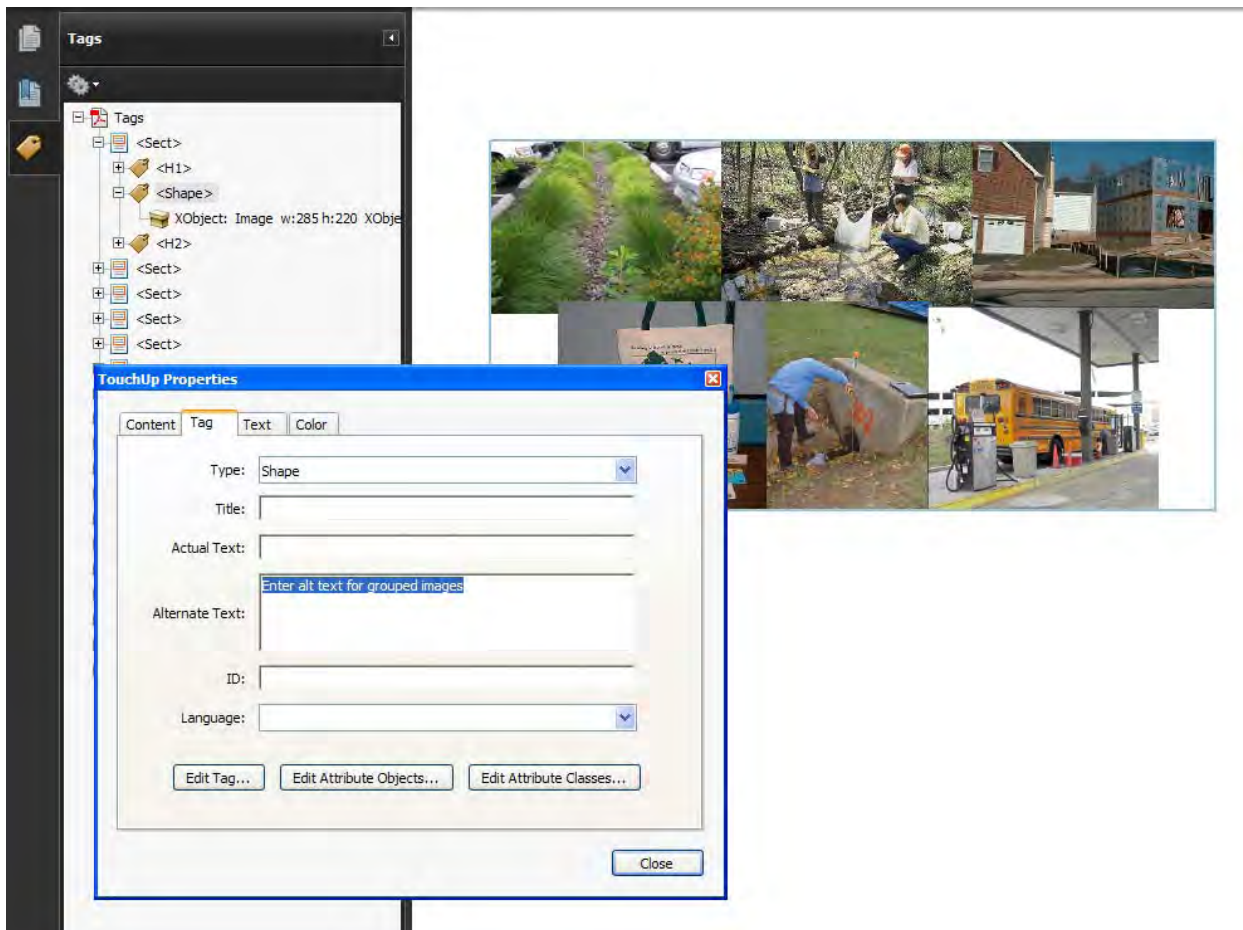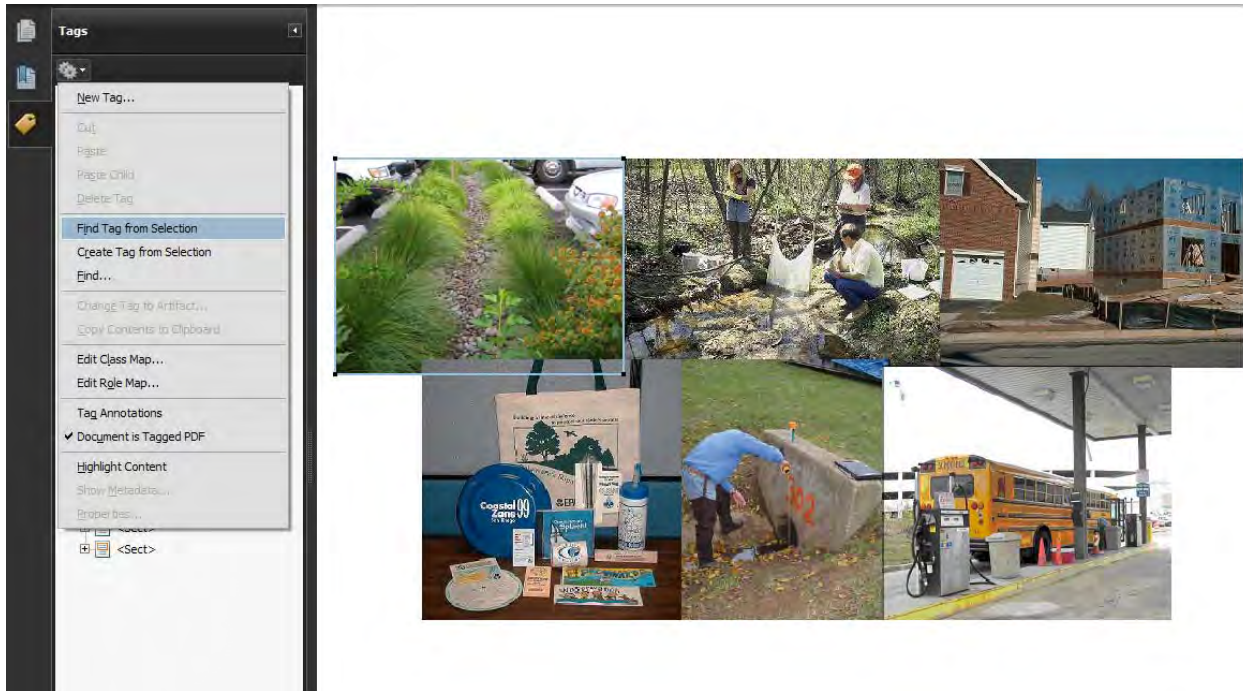
Note: For documents originating in MS Office products (i.e. Word, Excel, Powerpoint, etc…), it is much easier to add alternative text in the original document prior to converting to PDF by right clicking on the image. Choose Format Picture, then go to the Web tab and add alternative text.

5. **Run a Full Check Again**: Repeat step 2 and see what you still need to fix.

Why it's sometimes hard to make a PDF accessible

A) The document was scanned to PDF

You'll have to OCR it, because it literally is a picture of the page. OCR stands for Optical Character Recognition, and Adobe Acrobat has a built-in OCR feature. When it's done, your document will have a text layer that screen readers can access.

Note: If your document has a graphic that contains text, the Acrobat tagging utility cannot recognize that text, so you'll need to find the graphic in the tagging structure of the document and give it alt text.

To apply OCR Text Recognition Choose: *Document > OCR Text Recognition > Recognize Text Using OCR*
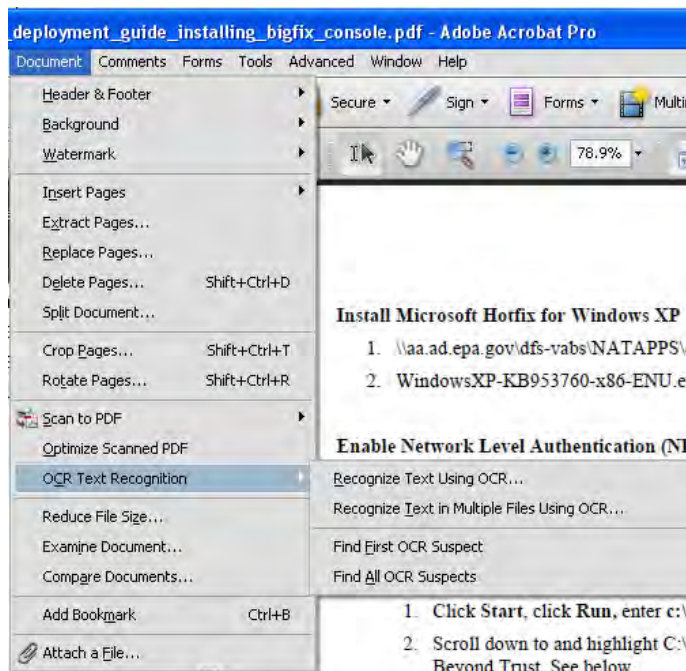
B) Reading order is messed up

You can repair it, or you can go back to the source document and take steps to prevent reading order problems in the first place.

Reading order problems may occur when the document contains complex formatting features such as multiple columns, text boxes, tables and text within graphics.
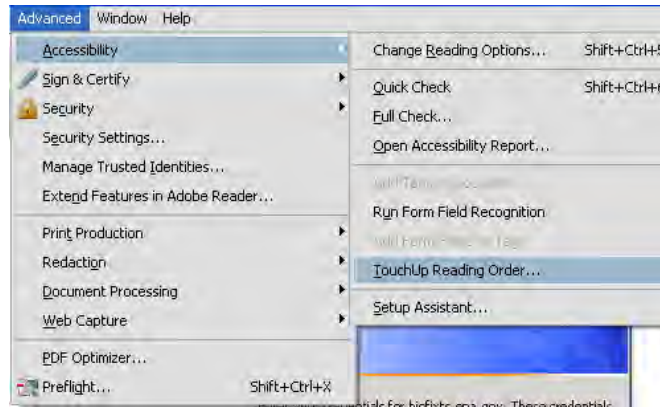
We often lay out documents with columns and text boxes and tables so that they'll be visually appealing when printed. When Acrobat tags a PDF with these features, the resulting code doesn't necessarily follow the same reading order that appears on the page. While you may look at a document and read it top to bottom, column by column, the code in a tagged PDF may skip all over the page. (See Appendix A for an illustration of jumbled reading order.)

C) Long complex documents

You can tag these successfully, too, but it's going to take time and patience. Issues can occur in small files as well as large.

Listed below are the steps that will help make a PDF file accessible.

1. Open navigational panels (*View > Navigational Panels*)
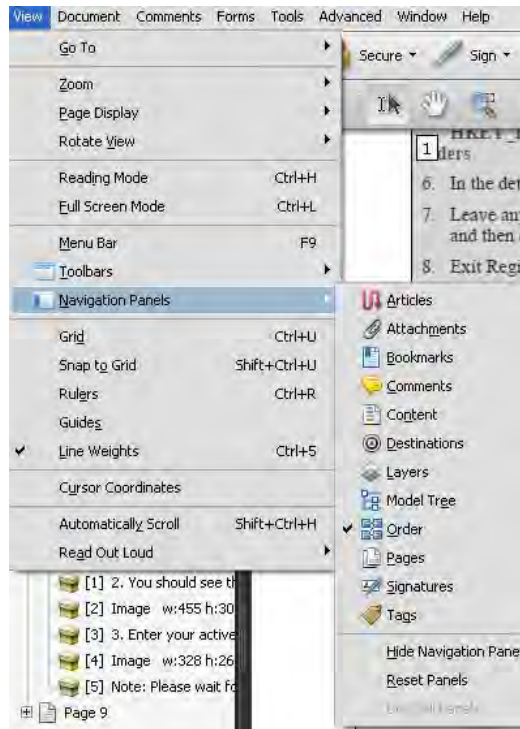
2. Choose the following navigation panels
   a. Bookmarks

> b. Content
> c. Order
> d. Pages
> e. Tags

These panels only have to be chosen once, they will appear in the left navigation each time you open Acrobat. Once these Navigation Panels are in place, you can begin making the PDF fully accessible.

3. Click on the header of each panel as they are opened; drag them (by the name of the panel) to the left edge of the page so that the icon of each panel can be viewed as an icon. As stated above, once these panel icons are present they will remain on subsequent openings of Acrobat until the icons are removed by the user. Once these Navigation Panels are in place, work can begin making the PDF fully accessible.
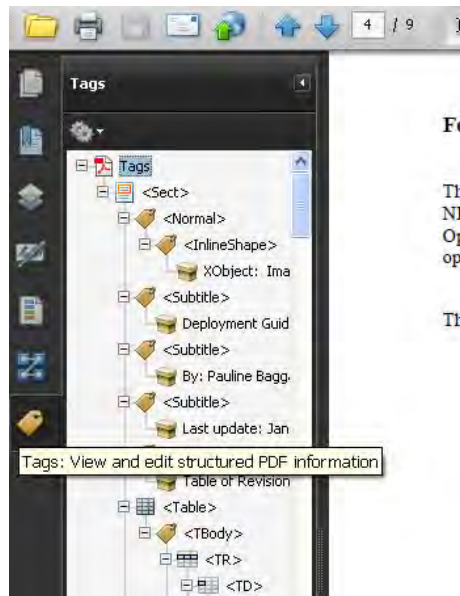
4. Below are a few errors that may come up when you check a document for compliancy:
    a. None of the elements or comments are contained in the structure tree
    b. Tab order may be inconsistent with structure order
    c. Elements with alternate text but no page content

In order to correct the error "None of the elements are contained in the structure tree:
- Click "Tag" icon located on the left side of the screen.
- Click "options"
- Then click "find"
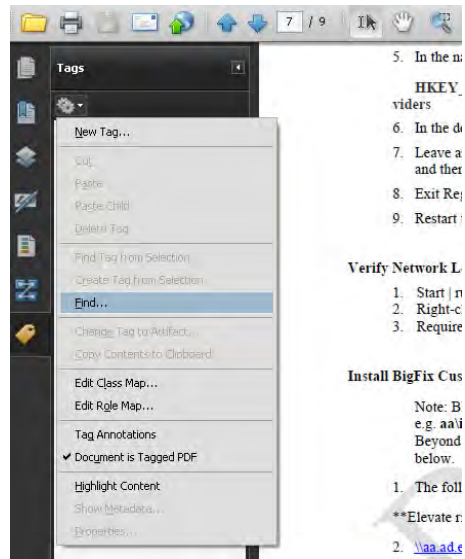
Start by choosing unmarked comments, choose radio button "search document" then click find. If nothing is found continue to the next choice in the drop down, continue this until unmarked annotations has been searched.
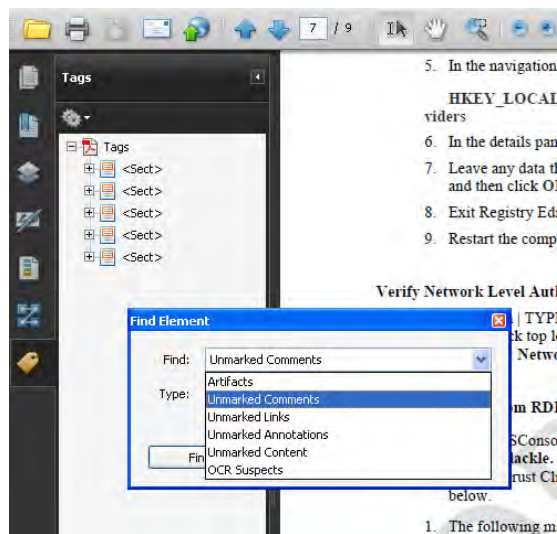
Once an error has been found, click on **Tag Element**.
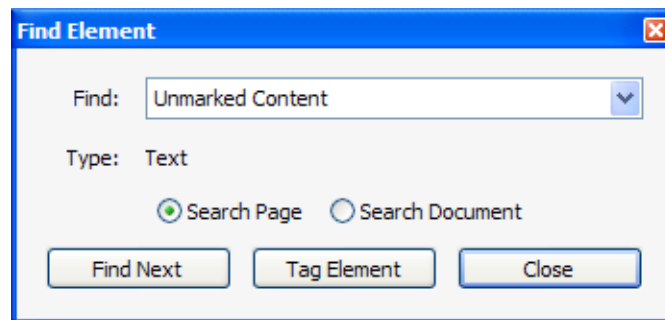
Then select the type of tag you want to create and click **OK**.
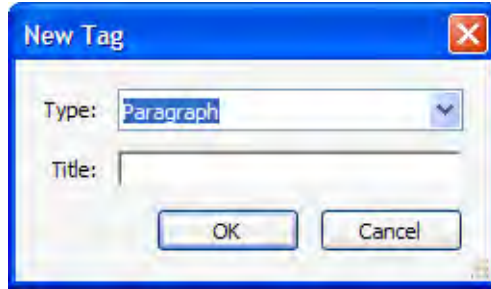
Once the error has been found again, click on tag element. This will tag the element that was causing the error. Repeat process until all errors have been corrected. Once this has been completed, run another full check on the document.

In order to correct the error "Tab order may be inconsistent with structure order" open the pages icon on the left. Click on the top page once with the left mouse button, scroll down to the last page of the document. Press shift + left mouse button, this will highlight all pages in the document.
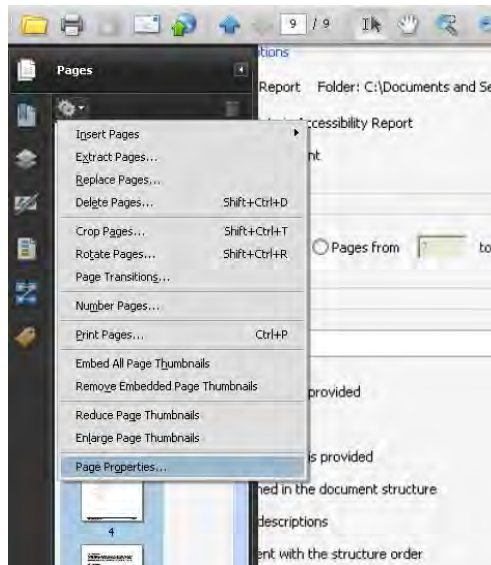
Click Options located at the top of the pages that have been highlighted. Choose page properties located at the bottom of the dropdown. Choose the radio button "Use document structure" then click ok.
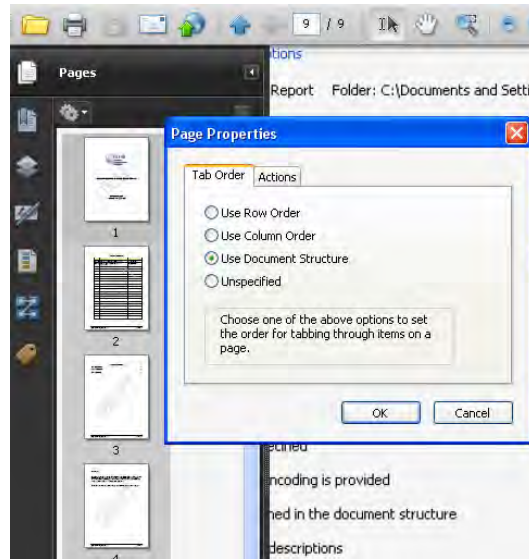
Once this has been completed, run another full check.

In order to correct "Elements with alternate text but no page content" open the content icon on the left. Use the ctrl key and click on the + sign to expand all tag layers of the file.
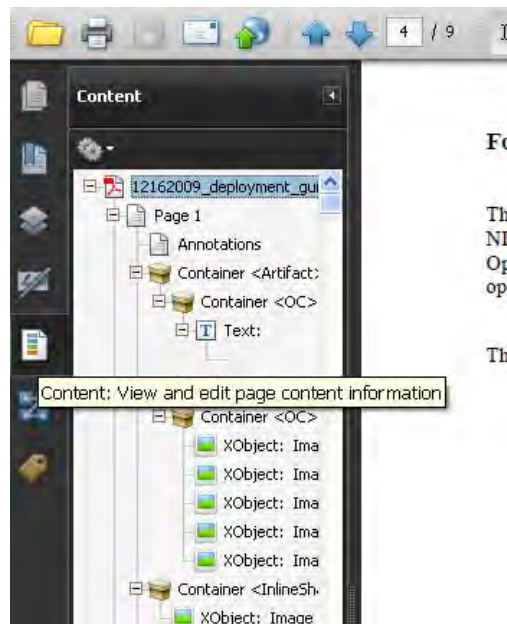
Look for tags that have nothing attached to them (for example a <p> tag that has no content within it should be deleted). Before you delete any tag, right-click on the tag and choose Highlight Content. If no content is highlighted within the document and the tag is empty, remove the tag. You can do this by using the Delete key.

What is a "reasonable effort" to make my PDF file accessible?

This is the single most time-consuming question that the Web Workgroup (WWG) PDF Subcommittee considered over three years. In the end, **it comes down to a personal judgment call** on the part of the person paying for the work. (If you are a federal employee and you are doing the work yourself, that person is you or your

supervisor who is concerned about how you spend your time.) This workgroup cannot formulate an answer that will work well for everyone, but we can suggest things to consider in making your judgment.

1. Is it very hard to make this PDF fully accessible? If not, then maybe you should go ahead and do that. A small effort on your part could make a big difference to someone else.

2. Is this document unusually important or popular? If so, it may be reasonable to spend more effort on making it fully accessible.

3. Is it critical to get this information posted immediately? If so, you may be able to justify posting a non-accessible version temporarily to have it up quickly, but if it is critical, it's very important to replace it with an accessible version later.

4. How much staff time or contractor money you have to spend will certainly figure into your decision.

If you decide to post a PDF that is not fully accessible, how do you comply with Section 508?

The following options were confirmed to satisfy Section 508 by the Section 508 workgroup of the CIO council in December of 2004:

1. **Provide contact information.** No one in the federal Web community seems to have much experience with this, so it's hard to predict what people will request. You'll probably have to handle this on a case by case basis, and work with your audience to give them what they need. Contact info only has to be provided once per page if it applies for all the PDF documents. It can also be provided in-line with the PDF document or in the disclaimer box.

    What kind of contact information should we provide, and who should we list?[1] We think it should be left up to your office. There is no "one size fits all" solution for this. Phone numbers and/or email addresses are good choices. Possible contacts include your library, public information center, public affairs office or the person who created the document. Whatever information you provide, the document should be traceable back to the program that produced it. See Section 4 - Linking, for several ways to provide this information.

2. **Don't use PDF.** Use HTML. If you must preserve the official pagination and structure of the document, or the graphics, or the signature, this option likely will not work for you. But for some documents it may be something to consider. Each Office and Region will need to decide whether they can offer documents only in HTML format.

Section 4 - Linking to PDFs[2] on HTML pages

The basic principle is: give people enough information so that they know what they're getting into when they click that link to your PDF. See the EPA Standard for linking to PDF documents:
http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-linking

---

[1] Doug Wakefield (retired in 2005 from the Access Board), suggested that the contact to provide help on a PDF doesn't need to be the author or someone very familiar with the document. He thinks most requestors would probably only need a printed copy that they can run through a reader and study at their leisure; even someone who isn't familiar with the document could provide that. (We don't know whether most vision impaired people would be able to print the document themselves.) If they want to discuss the content of a document that's inherently visual (a graph or an aerial photograph), maybe it would be best if they could speak or correspond by e-mail with the person who created the image. This is one reason why it's important to include source information in a PDF's Document Description.

[2] This section has been revised to reflect the PDF linking standard adopted by the EPA Web Council on July 12, 2006.

The following information is to provide more detailed information about the requirements than is contained in the standard.

Required information

The information you must supply whenever you link to a PDF:

- title of the document
- PDF
- number of pages
- file size
- a link to the About PDF page must be on the page using a disclaimer or the in-line link. Only one per page is required.

Here's an example:

> How they get those ships inside the bottles (PDF) (24pp, 240K, About PDF)
>
> Other information that you will provide sometimes

- document date (optional but recommended)
- EPA publication number (if applicable)
- dimensions for large pages (e.g. for a poster)
- contact information

Let's go through the different elements of linking to PDFs.

Link text

Include the title of the document[3] and the acronym "PDF" in parentheses in the link text. Our testing with visually impaired users showed that they are able to determine that a file is in the PDF format when links are formatted this way. It also helps users who have no visual impairment.

If you're in doubt, read the link text aloud and ask yourself, "Would I know enough about this file if that were all I could hear?"

Tell them how big the file is

After the link, show the number of pages and file size as unlinked text, separated by commas, in parentheses.

For the number of pages, abbreviate using "pp." If there is only 1 page use "pg".

The standard abbreviation for megabytes is MB. The standard abbreviation for kilobytes is K. For file size, use "K" for files smaller than 1000K and "MB" for files larger than 1000K. For MB, use one decimal place. It's easier to understand 1.5MB than 1503K. The goal is to give people a general idea how big the file is, so you don't need to use a lot of decimal places

---

[3] Here's an example of a situation where common sense suggests an exception: a set of documents that are all related to rulemaking and are grouped together in a list. Usually these documents (the fact sheet, the rule itself, the Regulatory Impact Analysis, etc., all have the same or similar title as the rule, so it seems quite redundant to list the titles -- **instead we use a description of the document as the link** (e.g., Fact sheet, Regulatory Impact Analysis, etc.).

Document Dimensions

Provide as much description for the reader as possible. If the PDF is a poster, for example, you might provide the original dimensions:

How Far From the Nearest Road? (PDF) (poster 36 x 96 inches, 1 pg., 3 MB)

Contact Information

If you are posting a PDF that you know is not 508 compliant, and you are not posting the same content in some other accessible format (e.g., HTML or text), provide contact information. Add this in-line at the file location or it can be part of the disclaimer text (in the box or paragraph). (If you supply an email address, follow the EPA standard for linking to email addresses: http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/emailaddresses)

Examples:

> If you have difficulty accessing these PDF documents, please contact Chris Robinson (robinson.chris@epa.gov) at 202-123-4576 for assistance.

> If you need assistance in obtaining the content of this document, please contact EPA's Water Permits Division at 202-564-9545.

Link to EPA's About PDF page

On every page which includes a PDF, you must link to EPA's PDF page (http://www.epa.gov/epahome/pdf.html), which explains what PDFs are and contains a link to the Adobe Web site where people may download the free Acrobat Reader. Only one link to EPA's PDF page is needed per HTML page.

There are two ways to do this.

The long way - disclaimer box or paragraph

Provide the following disclaimer text (or similar language) in a box or paragraph.

> You will need the free Adobe Reader to view some of the files on this page. See EPA's PDF page to learn more.

Here are examples from pages that use disclaimer boxes.

Figure 1: Example of PDF disclaimer using the tag <p class="disclaimer">
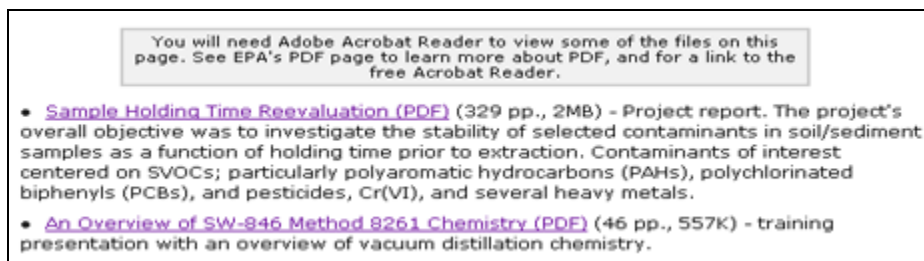
Figure 2: Example of PDF disclaimer using the tag <p class="disclaimer-wide">

Whether box or paragraph, the long disclaimer may appear:

- at the top of a page that includes many PDFs,
- right before a paragraph or list of PDFs, or
- immediately before the first reference to a PDF

Don't put the disclaimer at the bottom of the page or anywhere below the PDF links because people won't see it there.

- The disclaimer text may be in a box but it's not a requirement.
- The style sheet for the EPA Template includes styles for PDF disclaimer boxes. If you use a box for long disclaimers, use either the right-aligned box (Figure 1) or the centered box (Figure 2). `<p class="disclaimer">You will need Adobe Acrobat(...)</p>` or `<p class="disclaimer-wide">You will need Adobe Acrobat(...)</p>`
- You may include contact information in the same box.

When would you use this? When you have a lot of PDFs on one page.

The compact way – in-line disclaimer (About PDF link)

You may link to the "About PDF" page by adding "About PDF" as the last item in parentheses after the link to the document itself. Here is an example from a page that uses an inline disclaimer.

When would you use this? Use it when the boxed disclaimers are problematic, such as on a home page. When the box would be problematic to the design of the page, use a paragraph or inline disclaimer.

Other best practices for linking to PDFs

- Make the file size information in parentheses appear in a smaller font; this helps the title stand out visually. The style sheet for the EPA Template includes a style for PDF file information. `<span class="fileinfo">(114 pp, 611K)</span>`
- Include a title attribute `<a href="filename.pdf" title=" PDF, Document Title">` At the time of writing, screen readers can be set to read the title attribute, but that is not their default setup. That's why we consider using the title attribute helpful but not required.
- It's considerate to include the date of the document. One good place for this is right after the size in parentheses.
- File naming conventions:

- o Use all lower case letters
- o Do not use spaces or special characters in file names. Use hyphens instead of spaces or underscores. (This is necessary for the EPA search engine; spaces and special characters can break your URLs.)
- o Try to make the file name meaningful. It is often returned as part of search results and a meaningful file name can be beneficial. (e.g., pdf_white_paper_draft_v1_040110.pdf)
- o Having a unique file name is helpful when reading web statistics; using the EPA publication number as the file name makes it unique and easy to identify the document without having to open it (e.g., 420s06045.pdf).

Linking to multiple file formats

The 2010 workgroup does not recommend posting multiple versions of the same content. Evaluate your content and post either PDF or HTML but not both. Choose the best format for the content. We don't want redundant content.

- HTML if:
  - o Short documents.
  - o Large audiences because there are fewer accessibility issues.
- PDF if:
  - o If you need to preserve formatting.
  - o Brochures, flyers, newsletters and posters are meant to be printed. So why create an HTML version?

If you have content in two formats, just link them as distinctly as possible while identifying format and size.

The usability testing we did with assistive technology users didn't show conclusively whether they had problems with this technique, but it did suggest that they understood we were offering multiple file formats.

If you have more than one non-HTML version, try this:

> Impact of Best Management Practices on Water Quality of Two Small Watersheds in Indiana: Role of Spatial Scale (PDF) (120 pp, 1.02 MB) | Text Version (textfile) (154 K)

This one uses an inline About PDF link:

> Impact of Best Management Practices on Water Quality of Two Small Watersheds in Indiana: Role of Spatial Scale (PDF) 120 pp, 1.02 MB About PDF) | Word Version (MS-Word) (120pp, 540 K)

Section 5 - Making large documents more user-friendly

Chunking, dividing large documents into smaller parts, is not required or recommended at EPA.

Information on how to divide a large document into smaller parts has been provided in Appendix F. This process is called "chunking".   Should you decide that your audience demands a "chunked" document, there is a standard that requires the addition of a cover page for each chunk.
See: http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2

Section 6 - Metadata

Guidelines for adding metadata to Microsoft Office Products (Word, Excel, Powerpoint)

Metadata should be added to the Microsoft Office products before a document is converted to PDF.
See Appendix C.

Guidelines for adding metadata to your PDFs

When using metadata it is important that you apply it uniformly to your document collection and that you use it accurately. You are required to add metadata to PDF documents. To do this well, follow the guidelines on the EPA metadata site: http://yosemite.epa.gov/OEI/webguide.nsf/content/metadatafaqs

What is the Document Description? [4]

While viewing or editing a PDF, press Ctrl-D to see the Document Properties window (also found under File/Document Properties). Choose the Description tab view or edit the fields. (Note: You'll also hear people call this Document Summary; that's what it was called in earlier versions of Acrobat.) The Document Description contains four metadata fields: **Title**, **Author**, **Subject**, and **Keywords**.

Why is it important?

It is very important to complete the Document Description information in PDF files for four reasons.

1. Search engine optimization: your document is more likely to be included in search engine results.
2. Clearer search engine description: the correct title will be displayed in search engine results.
3. Document source information: if the author field is filled in, users will be able to find out what office wrote the document if they are looking for contact information.
4. Improved accessibility: The World Wide Web Consortium (W3C) recommends the use of metadata as a Priority 2 Checkpoint for web accessibility.

How do search engines use PDF metadata?

EPA's search engine will use the title metadata for the title in its search results. If no title metadata is available, the search results will state "no title available." Likewise, major search engines such as Google and Yahoo use the title metadata for the title displayed in the search results. If there is no title metadata available, the search results will display something like no title or a manufactured title.

---

[4] Acrobat also has something it calls Document Metadata. You can view this feature Acrobat 9 by clicking on "Additional Metadata" from within the Document Description window. All the information you enter in the Document Description fields transfers to the Metadata fields.

Document Description (Ctrl-D) if you fill this field	Advanced > Document Metadata the contents will show up in this field
Title	Document Title
Author	Author
Subject	Description
Keywords	Keywords

If you see a search result that looks like this C:\My Files\Policy\permits\faqs.doc, it means that it is in the metadata title field. The file owner did not accurately complete the metadata.

EPA's search engine will use the subject (or description) metadata for the document description in the search results. [5] If no subject metadata is available, the search results will display a system-generated description from the document text. Typically, the document description in the search results is acquired from the body text in external search engines.

Some public search engines will compare the keywords metadata to the document content to determine whether they are representative; if they are consistent, documents with keywords that contain search terms will get a higher score, if they are not, the document will be penalized for keyword spamming. Although not all search engines use the keywords, some do, so it is still worth completing the metadata for search engine optimization. The EPA search engine uses keywords metadata.

Completing the metadata for documents that are scanned directly into PDF is especially important if the text appears as an image instead of computer-readable text. Since the text of those documents is not searchable, the search engine will have to rely on the metadata.

Although search engines don't use the Author field, this information can be helpful to users who want to find out what office or program published the document. This can be particularly important to users who don't know or remember what web site it came from. If the user needs to contact the office for an accessible version, this might be the only place to look for the contact information.

Indexing PDF files

Indexing a document means that it has been added to EPA's collection and can be found via search by the public. Most new and updated documents will be indexed or re-indexed within 48 hours. If you have a large number of PDF documents that will be added, changed, deleted or moved, or if you have PDF documents that you know are not linked from any static pages, request a special crawl by sending an e-mail request to Susan Fagan or Judy Dew.

Check your PDF metadata reports

OEI has designed PDF metadata reports as a tool to help document owners improve the metadata in their PDF files. Each of the 4 metadata fields in the PDF document description is evaluated for quality. Overall, EPA does not have very good metadata so please review and update as needed.

These reports are available on the EPA PDF metadata page at: http://www.epa.gov/webmast1/metareports/pdf/.

Section 7 – Techniques for reducing PDF file size [6]

Reduce File Size

You can reduce file size by choosing **Document > Reduce File Size** but you will not have control over what Acrobat does to your file. The quality of your graphics might be reduced too much and embedded fonts might be removed.

---

[5] At this point, this feature is not working properly, but they're working on it.
[6] Note: Material in this section was adapted in part from Adobe's Acrobat Web tutorial.

PDF Optimizer

PDF Optimizer in Acrobat Professional gives options for reducing file size (***Advanced > PDF Optimizer***). In Acrobat Standard, the Reduce File Size command is the available for improving file size. You also have the option of reducing file size by using the Save As command. You can save the file under the same name or a different name. Know what you want before you take steps to reduce the size of a PDF file; consider how it will be used. If people will mostly view it online, then reducing the document's resolution is probably a good move. However, if you expect most people to print it and use the paper document (e.g. a map or pamphlet), then it may not print well if you reduce the document's resolution.

If your original document was created to be printed or sent to press, the graphics were probably set at high resolution. It is not ideal to use these files for a Web ready PDF. Once the original document is finalized, a copy of the document should be made to be used solely for creating a Web PDF. This way, the high resolution graphics can be replaced with lower resolution jpeg files or tiff files without affecting the quality of your original document.

Use RGB color or go grayscale

For Web-only PDFs, if you have a choice, use the RGB color space for your PDFs versus the CMYK color space. RGB has one less data channel than CMYK, so files are that much smaller in size. Also, Microsoft applications all think in RGB, even when importing CMYK images. CMYK is usually needed only for documents that will be professionally printed.

And if you can get away with it, grayscale or black and white images will save even more space.

To convert CMYK to RGB the process is as follows:

Within the PDF file choose *Advanced > Print Production > Convert Colors*
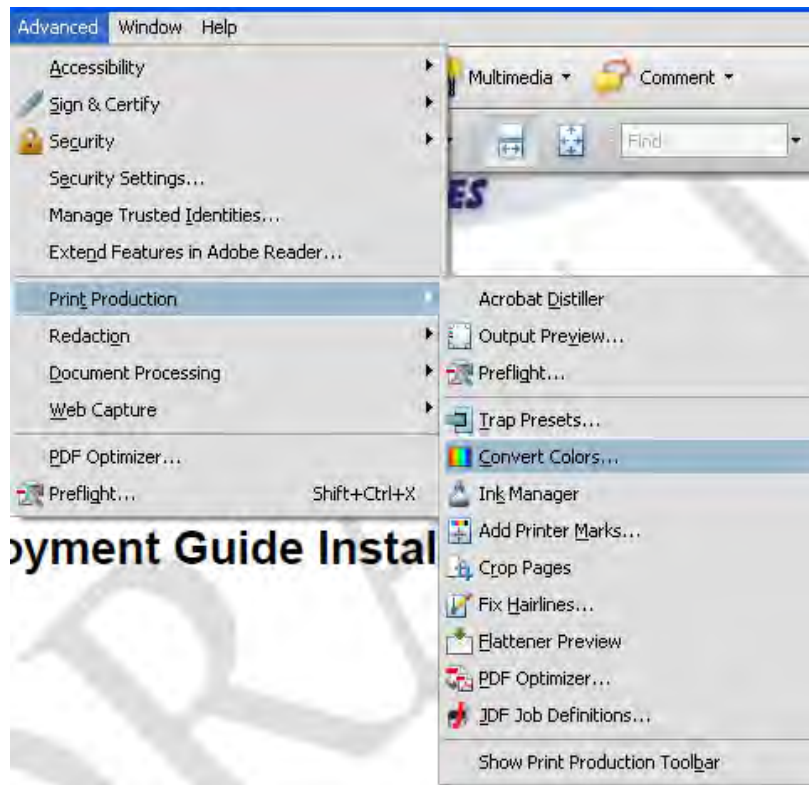
Once you are in the convert colors window, go to the matching criteria section and select image from the dropdown menu.
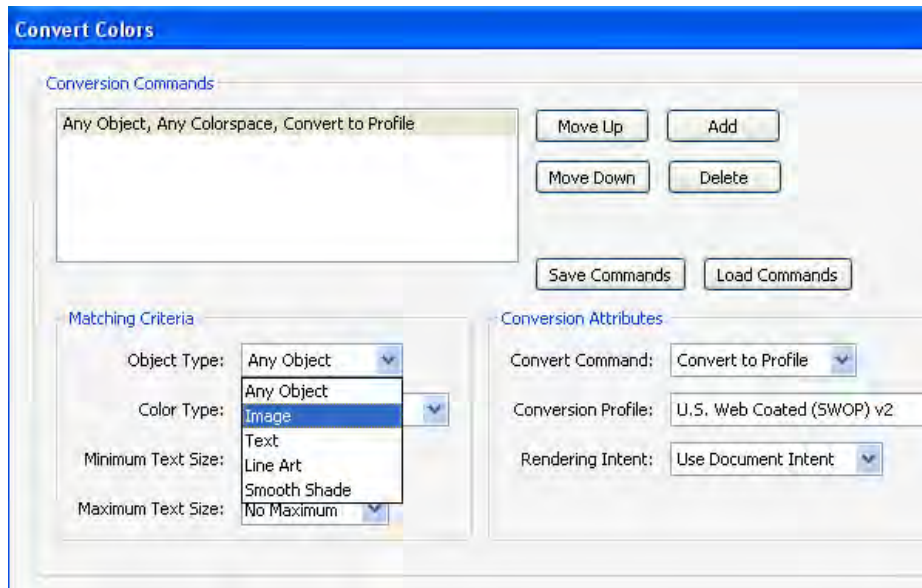
To Change CMYK images to RGB choose SRGB IEC61966-21.
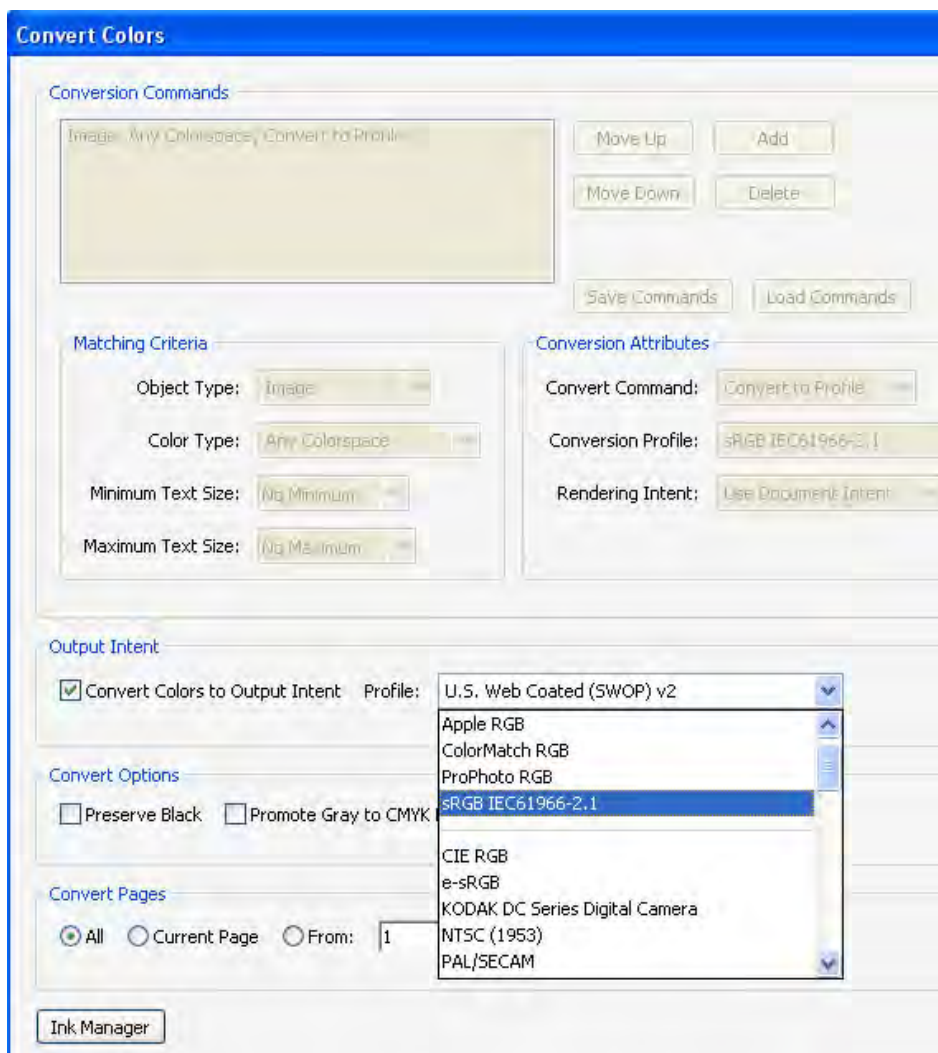
The next step is to check the box Convert Colors within the Output Intent section and choose Profile sRGB61966-21, Make sure that under the Convert Pages Section that the radio button "All" is checked.

To Convert to Grayscale (the process starts the same as converting to RGB.

Once in the Convert Colors window, go to the matching criteria section and choose image.

Within the conversion attributes choose Conversion profile Gray Gamma 2.2 from dropdown menu.
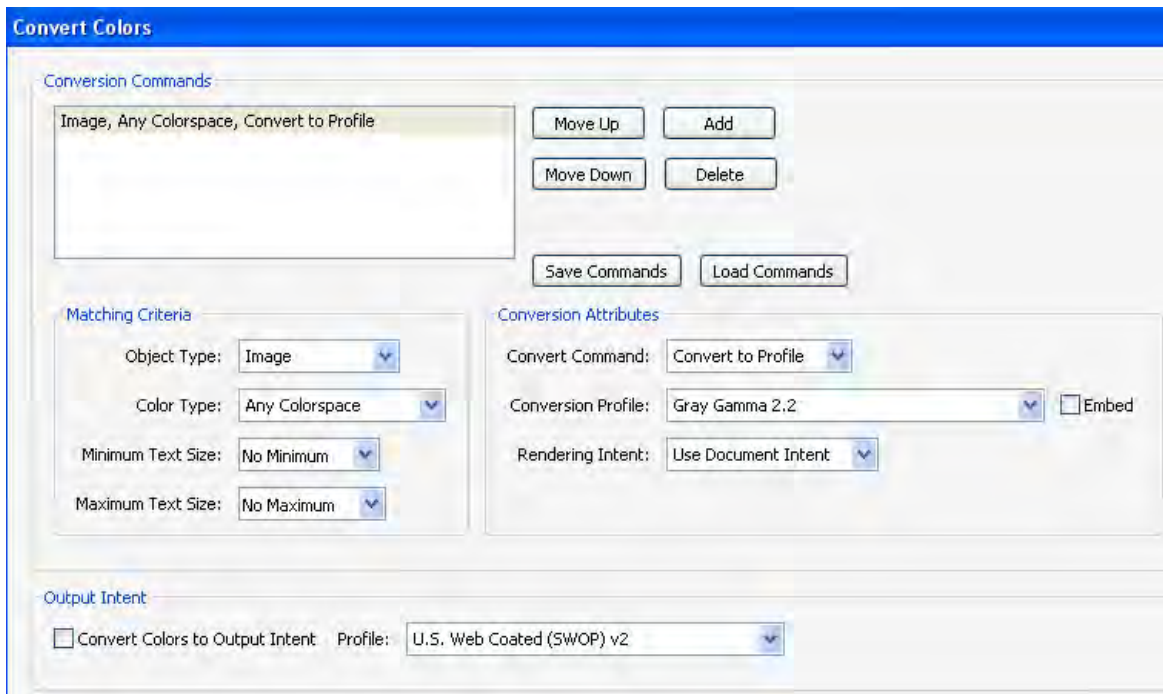
Next, check the box within output intent and choose profile Gray Gamma 2.2 from the dropdown menu.
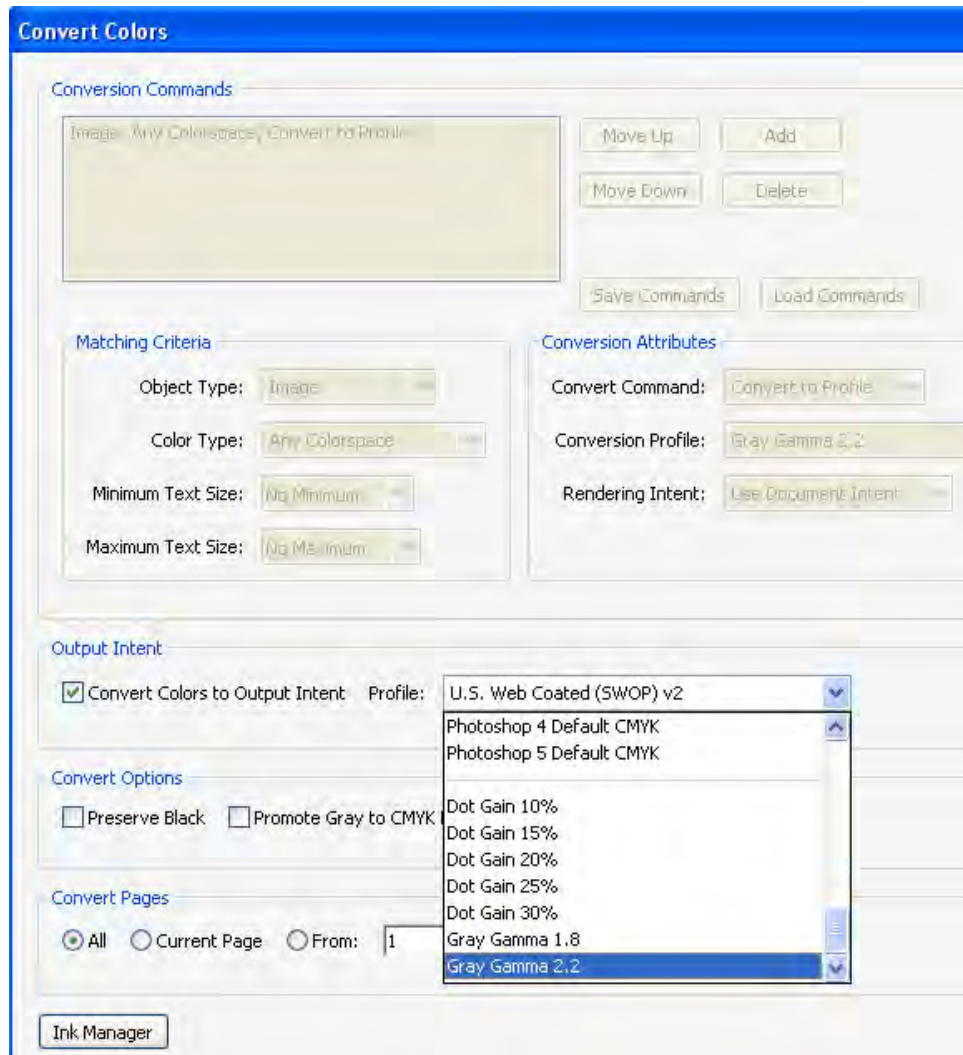
As with the process of converting to RGB, Please leave Convert Options unchecked and within convert pages section keep radio button selection to all.

Note: It may be best to convert images to RGB from CMYK or to grayscale within the Microsoft Office Product that the file is originating.

Use the "Save As" command

After you have made any final changes to the Adobe PDF document, choose **File > Save As**. Save the document with the same name to overwrite the original with your changes.

When you choose Save As, Acrobat rewrites the entire PDF document as efficiently as possible. When you choose Save, changes are appended to the final, which may increase the file size. By default, Acrobat also optimizes a PDF document for Fast Web View when you save it using the Save As command. Documents optimized in this way can be downloaded one page at a time from a Web server or network, reducing the time it takes to access and view them. To check and make sure that this option is selected, do the following:

- Acrobat 8 and 9: Go to **Edit > Preferences**. From the categories on the left, select Documents. Make sure that 'Save As optimizes for Fast Web View' is selected.

Audit the use of space in the file

The audit results can help you identify the most effective ways to reduce the file's size by showing which elements are significantly large. The audit reports the total number of bytes used by fonts, images, bookmarks, forms, and comments. It also shows the total file size. To view the audit:

- Acrobat 8 and 9: Choose **Advanced > PDF Optimizer**. Select the Audit Space Usage button. The audit results list the bytes used by each element and its percentage of the document's total size.

Optimize images

There are two ways of optimizing your images. The first is done in Acrobat after the PDF has been created, and the other is done in the distiller settings before creating the PDF.

(see following page for illustration)

In Acrobat

When using PDF Optimizer to optimize the images, it is a good idea to save the optimized PDF file with a new name. This way you can test different settings and leave the original PDF document unchanged. If you reduce the pixels per inch too low, the graphics may be unreadable. It is a good general rule not to go below 100 pixels/inch. To use the PDF Optimizer, do one of the following:

Choose *Advanced > PDF Optimizer*.

On the Images tab, select compression options for color, grayscale, and monochrome images.
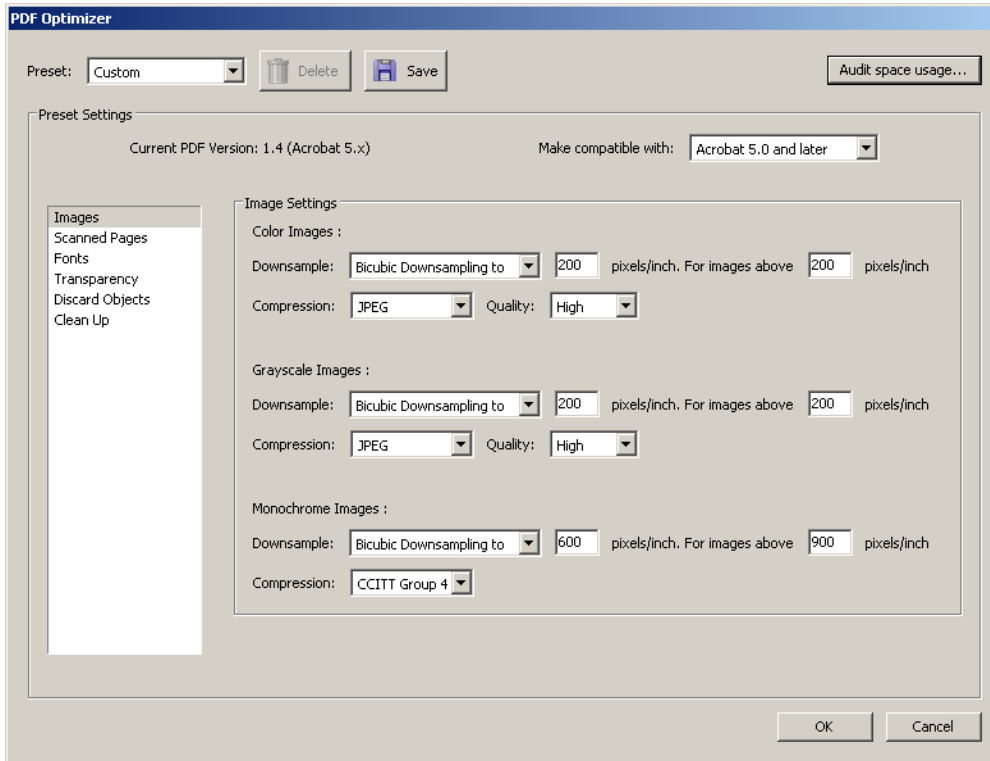
Downsampling lowers the resolution of images. Compression eliminates unnecessary pixel data. In general, you should use a compression setting of JPEG or JPEG2000 for photographs and other images in which color changes gradually. ZIP is good for illustrations with large areas of solid, flat color. JBIG2 works well for monochrome images. If you choose JPEG or JPEG2000 compression, select a quality level to determine how much pixel data is removed. Lossless (available only for JPEG2000) retains all pixel data.

In Distiller

If you use Acrobat Distiller to create your PDFs, it has similar options for image compression.

Distiller: Choose **Setting > Edit Adobe PDF settings**, then select the Image tag.
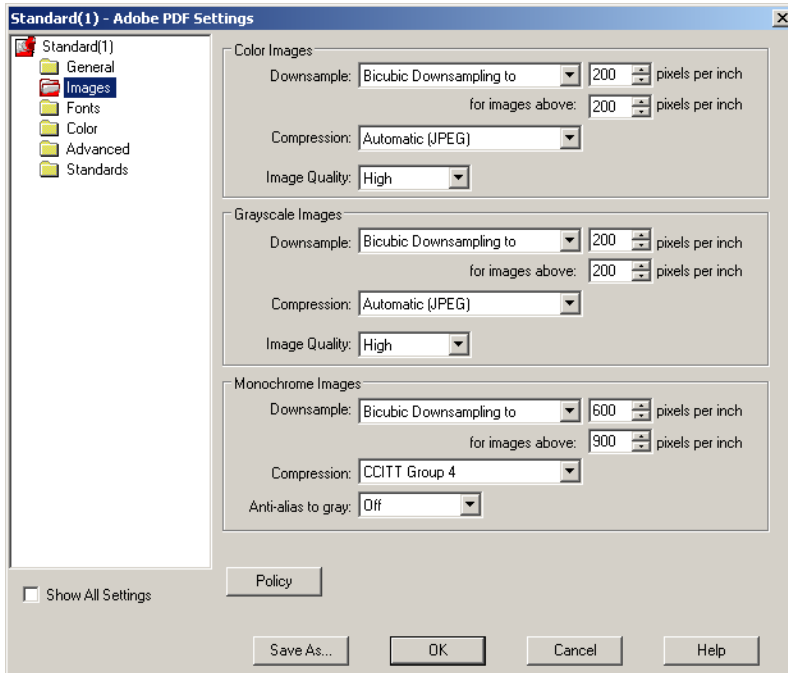
Embed only the fonts you need

Make sure you embed only fonts that are used in your document. Fonts often account for a large percentage of a PDF document's file size.

Distiller has an option that sets certain fonts to be always embedded. This can result in fonts being embedded that are not necessary. Acrobat allows you to remove them easily by choosing **Advanced > PDF Optimizer** and selecting the Fonts tab. Also, make sure you select the check box labeled "Subset embedded fonts when percent of characters used is less than: 100%".
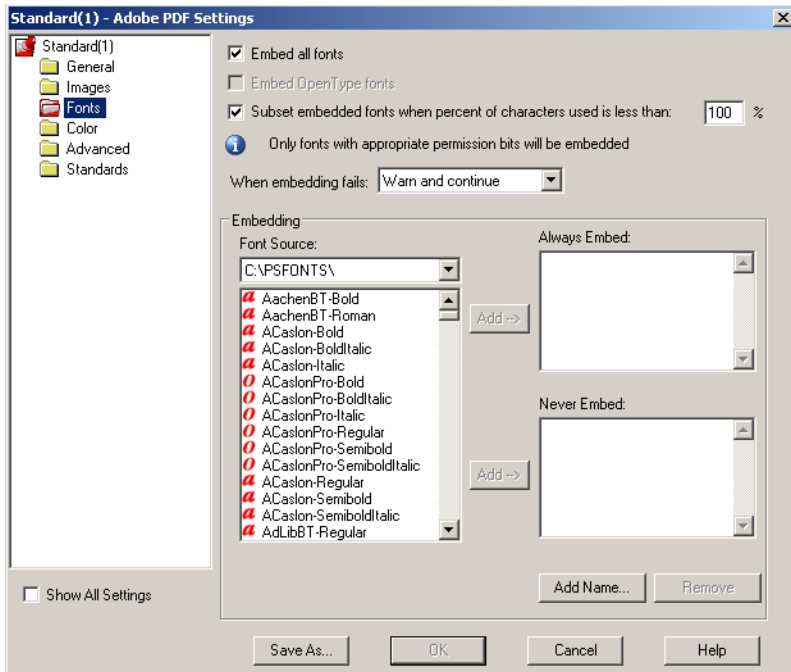
Acrobat Distiller: Choose **Settings > Edit Adobe PDF settings**, then select the Fonts tag.

Sometimes you can safely remove fonts, but do this only if you know that they are already installed on the computers of the people who will read your PDF document. If you remove a font that isn't available, Acrobat will pick a substitute font when the document is opened on their computer. Unless this is a basic font, such as Arial, Helvetica, Times, or Times New Roman, or similar basic fonts, the result can be unpredictable.

Try to minimize the number of fonts that you use in a document. Each font you use adds to the size of your document.

Remove unused elements from the document

Acrobat 9 has a Clean Up tab and a Discard Objects tab, which lets you remove elements from the PDF document that you don't need. The options that are selected by default do not affect functionality, but other options may. Use caution when selecting unfamiliar options. If you choose to experiment, save the optimized PDF file with a new name so that you leave the original PDF document unchanged.

Clean Up Settings

Choose *Advanced > PDF Optimizer* and select the Clean Up tab.
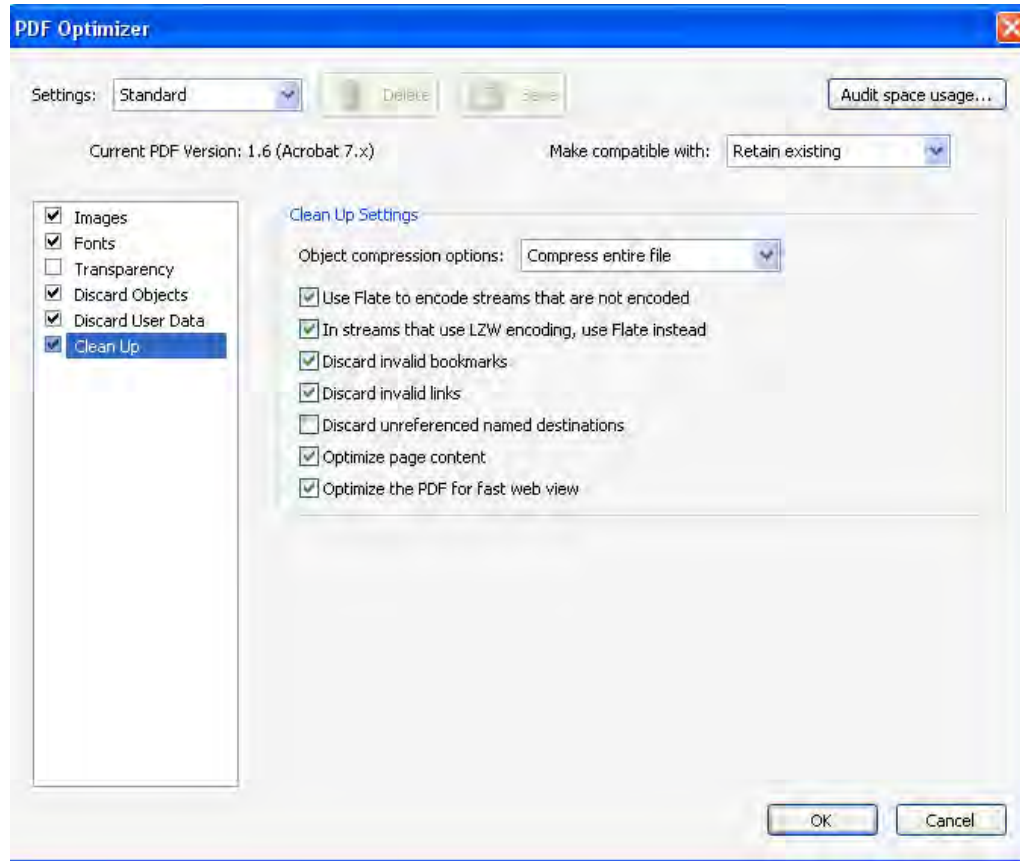
After selecting your options, select 'OK'.

Discard Objects Setting

Choose **Advanced > PDF Optimizer** and select the Discard Objects tab.
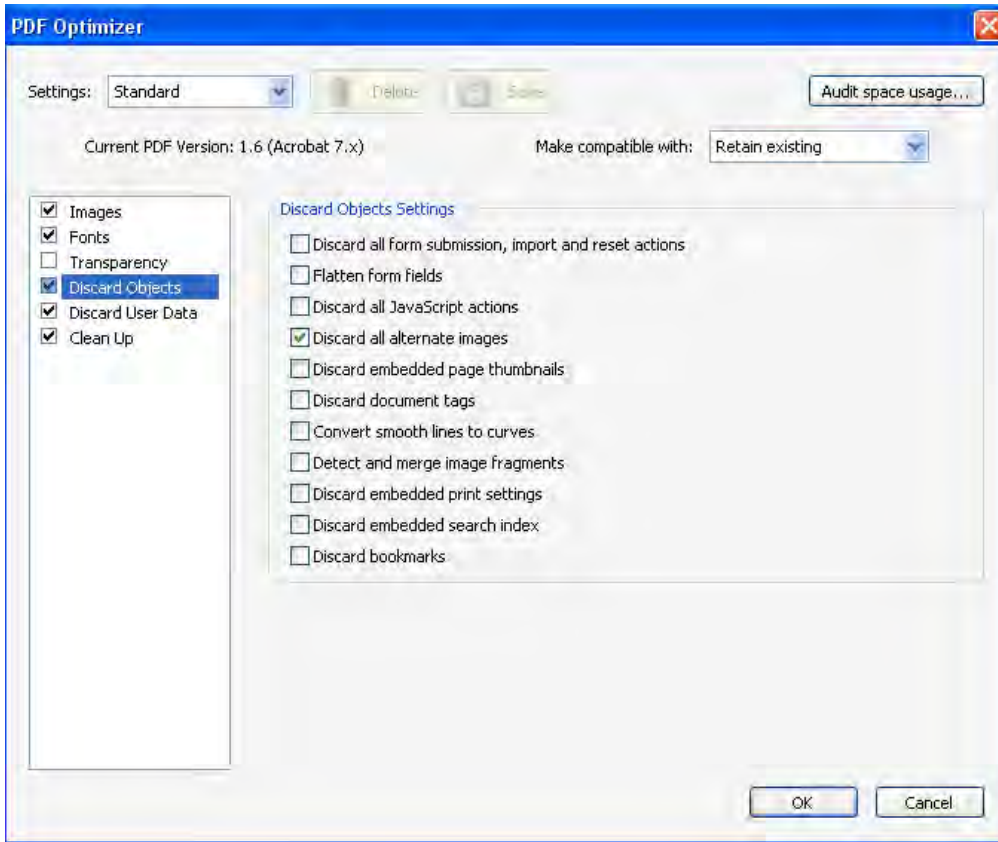
Discard Objects Settings 1

Warning: If your PDF has hyperlinks or is tagged, **do not** select the following check boxes:

- Discard all form submission, import, and resent actions
- Flatten form fields
- Discard document tags
- Discard bookmarks
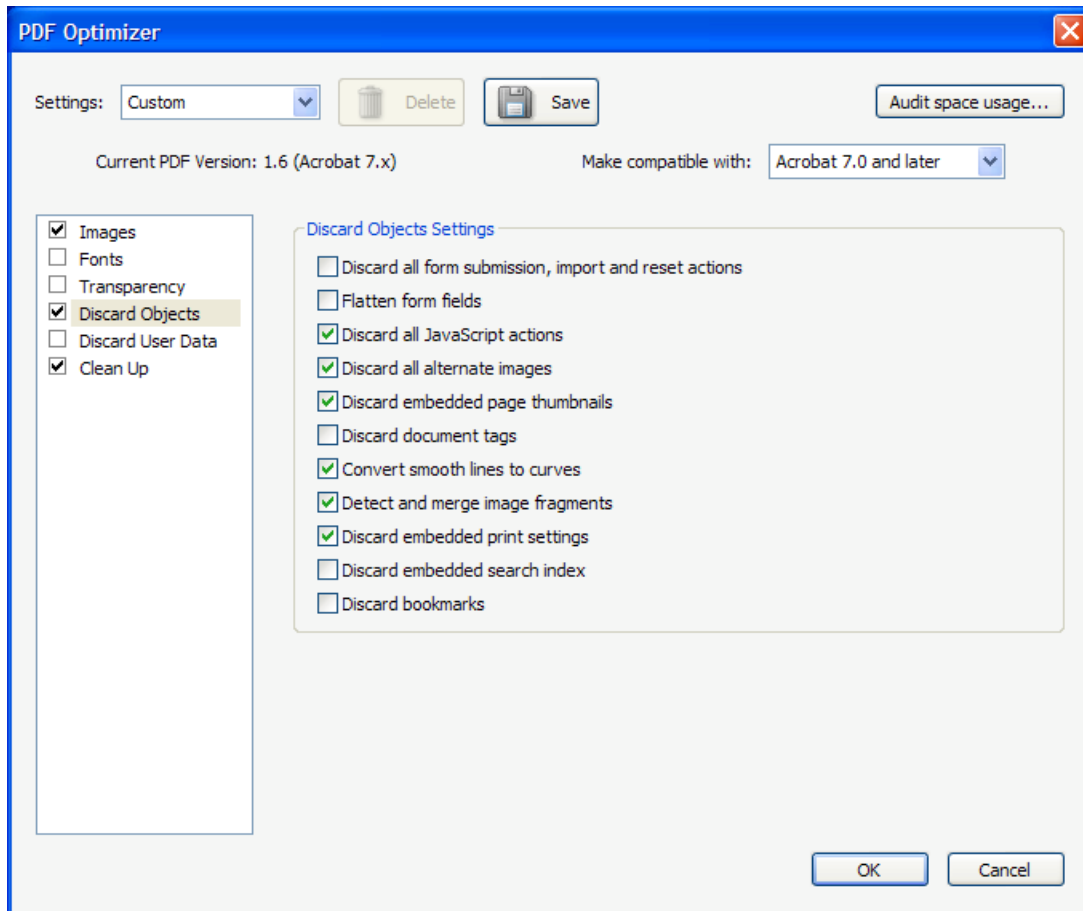
Section 8 – Dealing with scanned PDFs

General considerations

We don't recommend scanning paper documents to post on the Web. You'll get better accessibility results with electronic PDF conversion (Convert to Adobe PDF or Print to Adobe PDF). Scanned documents are images instead of text, so anyone using assistive technologies will not be able to "see" the document unless you use Paper Capture in Adobe Acrobat or some other Optical Character Recognition (OCR) software.
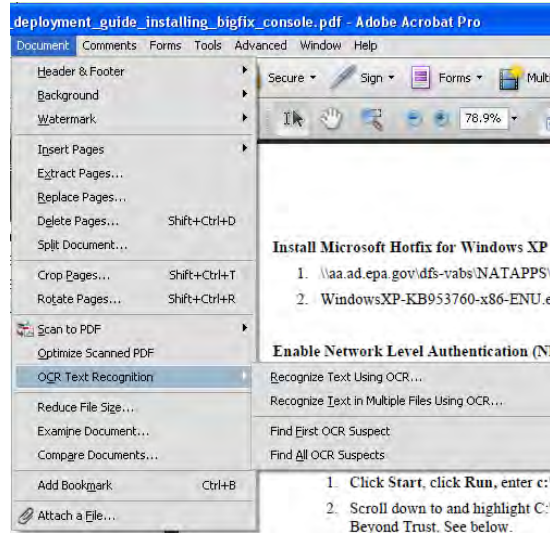
- If a document is older than July 2001, you do not need to do anything to it. If it is a frequently used file, we recommend that you make the material accessible, but that's not a requirement under Section 508.

- If it is a recently scanned document, you will need to post an accessible version or provide contact information so that people can contact you to request an accessible version as needed. (See Section 3 on Accessibility.)

What about using a scanned image of a person's handwritten signature?

We often make documents available electronically to satisfy the E-FOIA requirements[7]. Attorneys preparing for court are frequent requesters of EPA records and may need a copy of the signature page. Keep in mind that the web version of these documents would be courtesy copies, while the original with the signature is on file.

Our recommendations:

1. Don't scan a document to include a signature. Use /s/ or "original signed by "as a way to indicate a hard copy signature is on file. Ex: /s/ Charles Bert or Original signed by Charles Bert.

2. In cases where the record(s) has been requested 3 times or more or when the agency anticipates the documents may be requested frequently, we should **only scan the signature page.**

3. It's especially important to add metadata to a scanned PDF, particularly if you haven't used OCR on it. Since the text of scanned documents is not searchable, search engines will have to rely on the metadata you provide.

Section 9 – Which version of Adobe Reader should our PDFs be compatible with?

---

[7] Electronic FOIA Amendments of 1996: Three or more requests under FOIA for the same or substantially the same record qualify as a "frequent request".

We recommend saving PDFs to be compatible with Version 8 Readers. As we write this document, multiple versions of both the Adobe PDF creation software and the free reader are available. (See http://www.adobe.com/products/acrobat/acrrsystemreqs.html for complete details on the reader.) The readers are free downloads and have some backward compatibility. Due to the advances made in file compression and accessibility in newer versions of Adobe Acrobat Professional, we believe keeping the reader set to the current minus one version is a best practice. Another option is to set it to the current EPA Desktop Adobe Reader standard.

For example: Today Feb 17, 2010, the current version of the Adobe Reader is 9, but it's a recent release and most people do not have it deployed on their desktops. The EPA standard, today, is Adobe Reader 8. So save to Adobe Reader 8.

Section 10 – Does it matter which version of Acrobat I use to create PDFs?

We recommend getting the latest version of Adobe Acrobat Professional (currently Adobe Acrobat 9.0 Professional). Thankfully, Adobe has not changed the name of its PDF creation software; it remains Adobe Acrobat. However, there are now multiple editions of this product too: Professional, Standard, Elements, and Capture. The professional version has accessibility features and it is recommended; you can compare them at http://www.adobe.com/products/acrobat/matrix.html.

The 2010 workgroup, recommends keeping current with the Adobe Acrobat Professional version because the file compression and accessibility tools have improved with each release.

You can use other PDF creation software but it has to be able to create accessible documents and metadata that meets the Agency's standards.

Section 11 – Security features

Acrobat has security features you can use to prevent printing, copying, copying or extracting the content, and more. These can be defeated by people with readily available software and a little persistence. So while the security features may dissuade some people, all they really do is slow down a determined person who wants to defeat them.

Security restrictions can be a problem for legitimate users too, because if you need to go back and make changes to a document later on, you may not be able to if you don't have the password. This can happen when you inherit a document that someone else has created. Then you'll need to track down the password or the pre-PDF original, or hack the document yourself.

We don't recommend using security features to "lock down" some or all aspects of a document. If you have contractors preparing PDFs for you, specify that they should not use any document security restrictions.

Redacting PDFs is a different security issue. Some agencies have been embarrassed when they redacted documents in ways that could be defeated (e.g. masking words instead of deleting them).

You can read more about this in a paper published by the National Security Agency. There are also tools available for Microsoft Word that will remove change history from a document (so that people cannot see what edits were made). Please see Appendix D for these references.

Appendix A – PDF reading order problems – an illustration

1. This is what the page looks like.

Introducing...

The Nease Chemical site has a new remedial project manager. **Mary Logan** has spent almost 20 years with EPA in both Region 5 (Chicago) and Region 2 (New York). Mary has degrees in biology and environmental health sciences from the University of Chicago and University of Illinois respectively.

Contact EPA

If you have any questions, concerns, comments or want to be on the mailing list for news about the Nease Chemical cleanup, contact these EPA representatives:

Susan Pastor
EPA Community Involvement Coordinator
(312) 353-1325 or (800) 621-8431, weekdays 10 a.m. - 5:30 p.m.
pastor.susan@epa.gov

Mary Logan
EPA Remedial Project Manager
(312) 886-4699 or (800) 621-8431, weekdays 10 a.m. - 5:30 p.m.
logan.mary@epa.gov

As this project is being done in cooperation with Ohio EPA, questions can also be directed to:

Sheila Abraham
Site Coordinator
Division of Emergency and Remedial Response
Ohio EPA Northeast District Office
(330) 963-1290
sheila.abraham@epa.state.oh.us

2. When you try to select text, it doesn't highlight everything in the correct order.

Introducing...

The Nease Chemical site has a new remedial project manager. **Mary Logan** has spent almost 20 years with EPA in both Region 5 (Chicago) and Region 2 (New York). Mary has degrees in biology and environmental health sciences from the University of Chicago and University of Illinois respectively.

Contact EPA

If you have any questions, concerns, comments or want to be on the mailing list for news about the Nease Chemical cleanup, contact these EPA representatives:

Susan Pastor
EPA Community Involvement Coordinator
(312) 353-1325 or (800) 621-8431, weekdays 10 a.m. - 5:30 p.m.
pastor.susan@epa.gov

Mary Logan
EPA Remedial Project Manager
(312) 886-4699 or (800) 621-8431, weekdays 10 a.m. - 5:30 p.m.
logan.mary@epa.gov

As this project is being done in cooperation with Ohio EPA, questions can also be directed to:

Sheila Abraham
Site Coordinator
Division of Emergency and Remedial Response
Ohio EPA Northeast District Office
(330) 963-1290
sheila.abraham@epa.state.oh.us

3. This is what you hear:

"Introducing... The Nease Chemical site has a Mary Logan has spent almost 5 (Chicago) and Region 2 (New and environmental health sciences from the University of Chicago and If you have any questions, concerns, comments or want to be on the mailing list for news about the Nease Chemical cleanup, contact Susan Pastor Coordinator (312) 353-1325 or (800) 621-8431, weekdays 10 a.m. - 5:30 p.m. Mary Logan (312) 886-4699 or (800) 621-8431, weekdays 10 a.m. - 5:30 p.m. logan.mary@epa.gov As this project is being done questions can also be directed to: Sheila Abraham Site Coordinator Remedial Response (330) 963-1290 sheila.abraham@epa.state.oh.u s new remedial project manager. 20 years with..."

EPA United States Environmental Protection Agency

EPAXXX-X-XX-XXX
Month Year

Appendix B: Sample Cover Sheet

This is the Title
of the Full Document:
Including the Subtitle if there is one

This is the Title of the Section
or Chapter or Appendix

Name of Division
Name of Office
U.S. Environmental Protection Agency

You can use this space to add a disclaimer or contract information
if the document is associated with or produced from a contract.

Appendix C - How to Create Web-ready PDFs from Word, Excel, and PowerPoint files

Step 1: Fill out Document Properties (File → Properties)
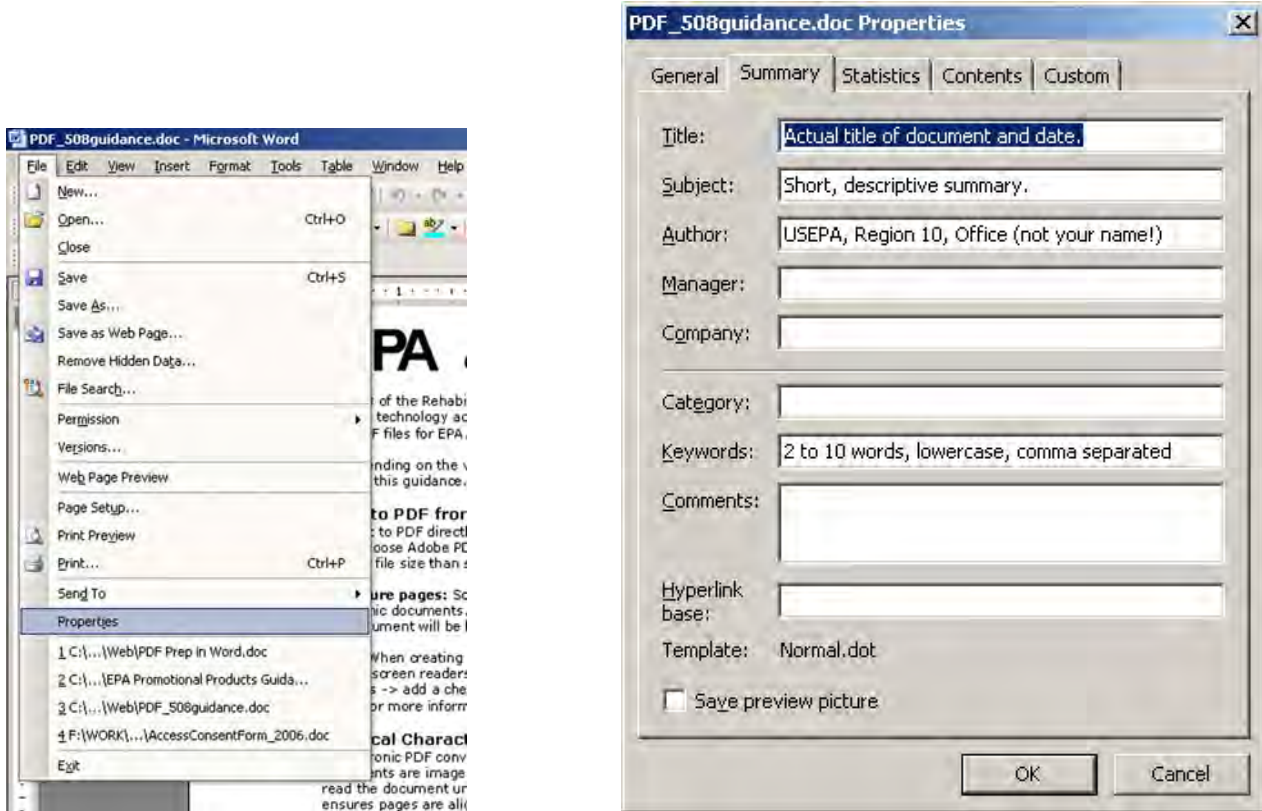
- follow the guidelines on the EPA metadata site: http://yosemite.epa.gov/OEI/webguide.nsf/content/metadatafaqs
- Complete title, subject, author and keywords fields

Step 2: Check Links to Websites

- All links to websites (URLs) in your document **must include http://** or they will not be converted into active links in your PDF (example: http://www.epa.gov)
- **Keep link text on one line**. Links that wrap to a second line will not be converted into active links in your PDF. (Note: If a URL wraps to a second line **and** is hyperlinked in MS Office, the hyperlink will convert correctly.)

Step 3: Adding Alternative Text to Images

Right click the image. Go to Format Picture. Go to Web Tab and fill in alternative text.

Step 4: Enable Document Tagging

Choose Adobe PDF → Change Conversion Settings. Be sure there is a check next to "Enable accessibility and reflow with tagged Adobe PDF". This helps ensure electronic screen readers will be able to recognize the document and read it aloud.
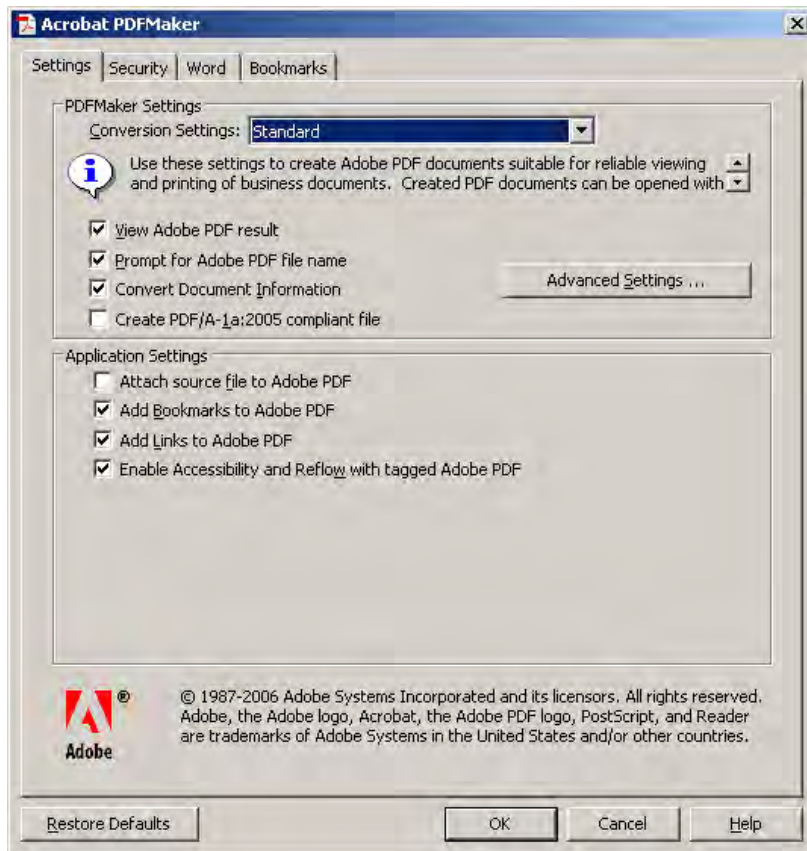
Step 4: Compress images to reduce file size

If your document or presentation contains photos, maps, or other images, this step will ensure your PDF file size is small for fast Web viewing.
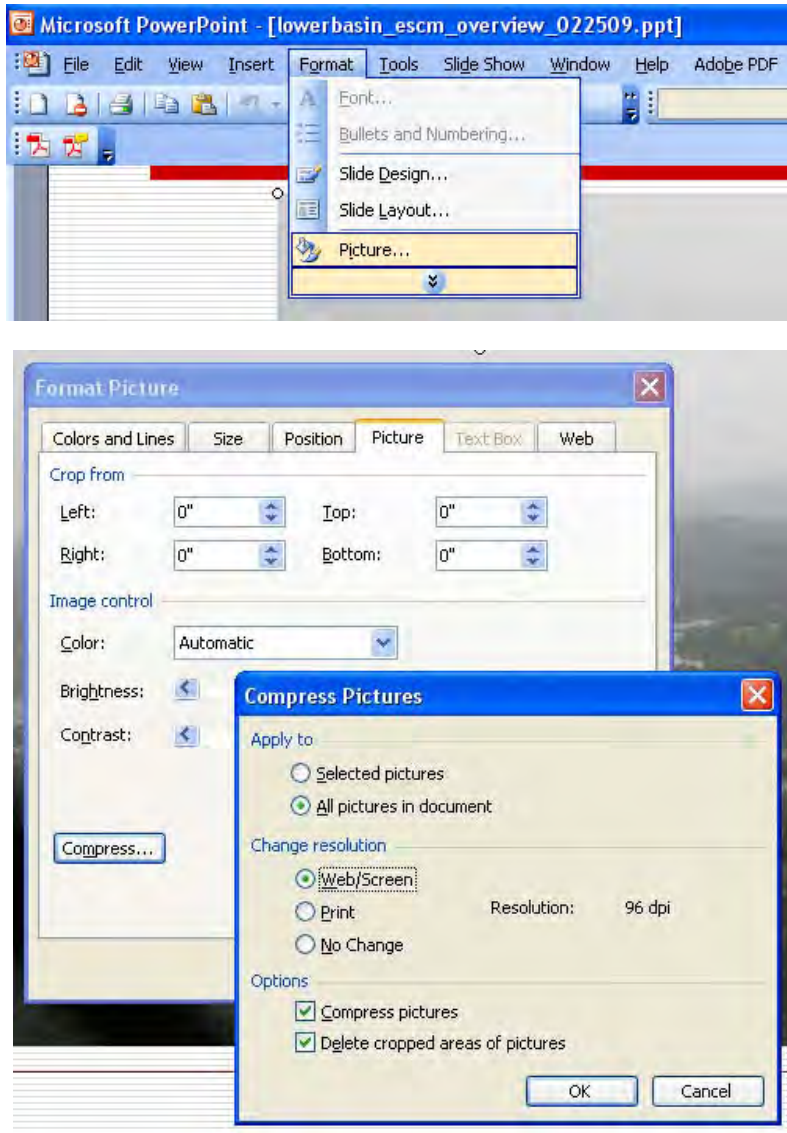
Click on any photo or other image in your document and then choose **Format → Picture**. Click on "Compress" in the Format Picture dialog box and another dialog box titled Compress Pictures will appear. Choose the following settings:

- Apply to: All pictures in document
- Change resolution: Web/screen
- Options: Be sure Compress pictures and Delete cropped areas of pictures are checked

Step 5: Convert to PDF (Adobe PDF → Convert to Adobe PDF)

Naming your file: When saving your PDF, choose a short file name and use hyphens and underscores rather than spaces. File names with spaces are illegal and creates problems on posted servers.

Appendix D —Checklist for Using PDFs on the Web

March 2010, EPA Web Workgroup PDF Subcommittee

Note: All directions assume you are using Adobe Acrobat 9 Professional.

1. **Determine whether PDF is the appropriate format to use** to post your information on the Web. Reasons to use PDF:
   - ❑ the document is longer than 5 pages, or
   - ❑ you need to preserve the formatting or layout of the original document (e.g. for printing), or
   - ❑ you need to publish a paper document for which no electronic version exists.

2. **If you are creating the PDF yourself, start with a "clean" original document.** Keep the format as simple as you can and use styling built into the word processor. This makes it easier to tag the document for accessibility.
   - ❑ Avoid placing graphics, text or tables side by side.
   - ❑ Use standard fonts.
   - ❑ Use styles, like Heading 1, Heading 2, etc.
   - ❑ Use the function for creating bulleted and numbered lists rather than formatting list items manually.
   - ❑ Use tabs and columns rather than spaces and tables for laying out your text.
   - ❑ Add alt text to graphics.

3. **Create a PDF using Convert to Adobe PDF.** Start with any authoring software (Word, Excel, etc.).
   If you are creating the PDF from an MS Office application, choose the Acrobat Tab located underneath the header (on the far right of the home tab) choose *Settings* and click the following boxes: Create Bookmarks, Add Links, and Enable Accessibility and Reflow with tagged Adobe PDF. If not available use: *Save as > PDF* (use default settings check box which include, open file after publishing and optimize for online and printing) Add file name and choose location the file should be stored.

   (Do not use *File > Print > Adobe PDF* that you used in previous versions of Acrobat unless the Adobe PDF menu is unavailable in your authoring software. You lose the document metadata and do not retain bookmarks if you use this method.)

4. **Create a Document Description**. Search engines use metadata from the Document Description in their search results.

   - ❑ Open the Document Description *File > Document Properties* (or Ctrl-D)
   - ❑ Select the "Description" tab.
   - ❑ follow the guidelines on the EPA metadata site:
     http://yosemite.epa.gov/OEI/webguide.nsf/content/metadatafaqs

5. **Create Bookmarks.** Select text and use *Ctrl-B* to insert a bookmark. (This method works after the PDF file has been created). To create bookmarks while the conversion to PDF is in progress, choose the Acrobat Tab located underneath the header (on the far right of the home tab) Choose *Preferences > Bookmarks > Convert Work Headings to Bookmarks* then select the OK button

6. **(Optional) Create Web links from URLs in text.** *Advanced > Document Processing > Create Links from URLs.* Create from URLs in document. Only URLs that begin with http:// will become links.

7. NOTE: EPA does not have a requirement to break large documents into sections. **If you decide to break a large document into sections** follow the requirements at:
   http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2

Accessibility Steps

1. **Tag Your Document**: (NOTE: You can skip this step if you converted your original document as a tagged PDF, e.g. in Word.)
   - If you are creating the PDF from a Microsoft Office application, choose **Adobe PDF > Change Conversion Settings** and select "Enable accessibility and reflow with Tagged PDF". Then select OK. Now create the PDF using **Adobe PDF > Convert to PDF** (or the Convert to PDF icon).
   - If you are using an existing PDF, tag the file for accessibility using Adobe Acrobat (**Advanced > Accessibility > Add Tags to Document**)
2. **Run a Full Check**: (**Advanced > Accessibility > Full Check**), then look at the report it generates and see what you need to fix.
3. **Decide whether to finish making the PDF accessible** or make the content accessible another way.
4. **If you decide not to make the PDF fully accessible,**
   - ❑ Provide contact information or
   - ❑ Use HMTL instead of PDF.
5. **If you *will* be making the PDF accessible, keep going:**
6. **Check the Reading Order**: Check the reading order and correct what doesn't make sense (**Advanced > Accessibility > Touch Up Reading Order**)
7. **Include Alt Text for Images**: Add alternate text for images (**View > Navigation Tabs > Tags**, select the image that needs alt text, then from the Tag window select **Options > Properties** and fill in the Alternate Text)
8. **Run a Full Check Again**: Repeat step 2 and see what you still need to fix.
9. **Fix any accessibility errors** identified. (see http://adobe.com/enterprise/accessibility/ for more info)

Follow EPA's PDF linking standards, posted at http://ww.epa.gov/webguide/create/link_standard.html

Appendix E - References

From EPA

EPA's "About Portable Document Format Files" page:
http://www.epa.gov/epahome/pdf.html

EPA's PDF Subcommittee on the intranet has many resources:
http://intranet.epa.gov/webgroup/committees/PDF/index.html

How to link to a specific page within a PDF
http://intranet.epa.gov/webgroup/resources/index.html#pdflink

PDF file summary problems in your TSSMS account:
http://www.epa.gov/webmast1/metareports/pdf/

From Adobe

See what Adobe has to say for itself and its PDF products
http://www.adobe.com/products/acrobat/main.html

Adobe Acrobat accessibility resources
http://www.adobe.com/enterprise/accessibility/main.html

Creating Accessible PDF Documents with Adobe Acrobat 7.0 (PDF, 115pp, 10MB)
http://www.adobe.com/enterprise/accessibility/pdfs/acro7_pg_ue.pdf

Other Resources

Planet PDF – the home of the PDF community
http://www.planetpdf.com/

Section 508.gov
http://www.section508.gov/

Facts and Opinions about PDF Accessibility by Joe Clark, August 22, 2005
http://www.alistapart.com/articles/pdf_accessibility

Steps toward Integrating Accessibility into Development of an Internet Option for the 2010 U.S. census
Elizabeth D. Murphy, October 2005
http://www.census.gov/srd/papers/pdf/rsm2005-11.pdf (PDF, 10pp, 161K)

Redacting with Confidence: How to Safely Publish Sanitized Reports Converted From Word to PDF
National Security Agency Report #I333-015R-2005, December 2005
http://www.fas.org/sgp/othergov/dod/nsa-redact.pdf (PDF, 14pp, 681K)

Federal Computer Week article about PDF/A
http://www.fcw.com/article89385-06-27-05-Print

Jakob Nielsen's Alertbox, August 29, 2005: Open New Windows for PDF and other Non-Web Documents
http://www.useit.com/alertbox/open_new_windows.html

Jakob Nielsen's Alertbox, July 14, 2003: PDF: Unfit for Human Consumption
http://www.useit.com/alertbox/20030714.html

Jakob Nielsen's Alertbox, June 10, 2001: Avoid PDF for On-Screen Reading
http://www.useit.com/alertbox/20010610.html

Appendix F – Breaking large documents into chunks

Why are large documents a problem?

Large documents pose two main kinds of problems:

A person may not want to download a really large file. They may not have a fast Internet connection. They may not be willing to spend the time it takes to get the whole document. They may only want part of that enormous document.

It can be easy to get lost in a long document. For most people, flipping to the page you want is easier in a paper document than in a PDF.

You may be able to reduce the file size of large documents (see Section 7 of this paper for techniques).

Add bookmarks within the PDF

Bookmarks can really help people navigate through a long document. If the bookmarks pane displays by default when the document opens, it lets people see the menu of bookmarks.

- Include bookmarks in any PDF ten pages or longer. (You can use them in shorter documents if you think they'll be helpful.)
- Follow the document's Table of Contents, if one exists; otherwise use the major headings of the PDF document.[8]

How to create bookmarks and make them visible
- Open bookmark navigation panel.
- Go to the first major heading you wish to bookmark
- Highlight the heading text
- Press *Ctrl + B*, this inserts the bookmark using the title that was highlighted
- Repeat process throughout the document
- Set the initial view to open with bookmarks visible (Under Document Properties - ctrl+D, Initial View, Navigation Tab: Bookmarks Panel and Page)

---

[8] For an example of a PDF with bookmarks based on the table of contents, see:
Town of Pines Administrative Order on Consent (PDF, 47pp, 3.3MB) April 2004
http://www.epa.gov/region5/sites/pines/pdfs/pines-groundwater-AOC-200404.pdf
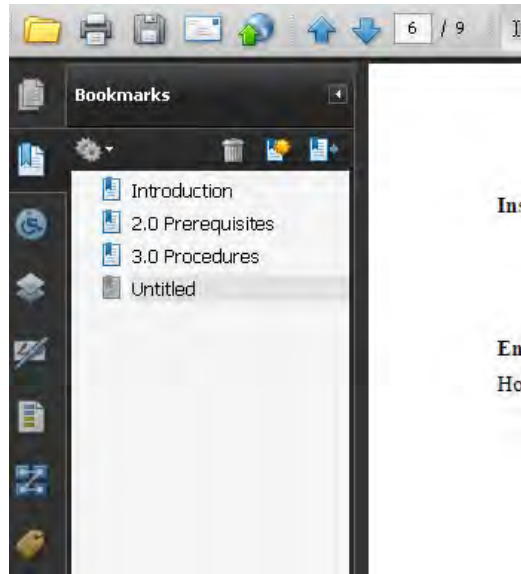
For detailed instructions on how to create and manage bookmarks, see the Using Bookmarks topic in Adobe Acrobat Help.

Provide an abstract or an executive summary

On the Web page that links to the document, provide a description (abstract, executive summary, blurb) that will help people learn about the document before downloading it. Here is an example of an abstract:

> Arsenic Removal from Drinking Water by Adsorptive Media (PDF, 1151 Kb, 55 pp)
> http://www.epa.gov/ORD/NRMRL/pubs/600r05159/600r05159.htm

Break up the document into chunks

Chunking, dividing large documents into smaller parts is not required or recommended at EPA.

The 2010 workgroup feels that at 50MB, you may consider the document content, the expected audience and make a determination if document chunking might be useful.

If a file size is > 50MB and you have determined to chunk a document to make it more usable, these are recommended best practice **options**:
- Put up an HTML page that contains the Title, Executive Summary (chunked document)
  o Include table of contents linking to individual chunks or link to a chunk that includes bookmarks to the other chunks.
- Post the document to an FTP site (complete document)
  o Link to it from an HTML page that contains the Title, Executive Summary
- Add information on how to order the document on DVD or CD. (complete document)

The choice of posting a chunked document on the Internet or providing complete documents on FTP or DVD/CD should be based on what can be well managed by the local web team.

For chunked documents:
- If a document is chunked, do not post the entire document just post the chunks (creates redundant content).
- Include required document info on each chunk (EPA Standard)

Note: The recommendation to consider chunking only applies to files above 50MB. It's not recommended that a document be chunked at all but you should consider if YOUR audience needs the document in smaller sections. There should be very few documents being chunked.

If you decide to break a PDF into sections, it is important to **provide a description (abstract, executive summary, blurb) on the Web page that links to the document chunks**. There are multiple ways to link to the chunked documents. Two are described below.

- Create a table of contents on the HTML page that links to each section. Each link should be accompanied by some description of that section. (Navigation on the HTML page)
  Example: http://www.epa.gov/nerleerd/westmethman.htm
- Link to the first chunk of the document that contains bookmarks linking to the other chunks. (Navigation contained in the PDF document.)
  Example: http://www.epa.gov/region1/superfund/sites/suttonbrook/273576.pdf

When you chunk a bookmarked PDF, create the bookmarks first, then break up the document using "Delete Pages" (i.e., snip away the bits that aren't going to be in this chunk and Save As chunk number whatever). This way you won't lose bookmarks. (Using "Extract Pages" may cause you to lose bookmarks.) However, you will wind up with superfluous bookmarks in the smaller chunks. You'll have to go through and delete those by hand.

Identify the chunks

This is a requirement for divided documents. See the EPA standard for each segment of a large document:
http://yosemite.epa.gov/OEI/webguide.nsf/standards-guidance/pdf-when2

Suppose a chunk of one of your long documents turns up in Google search results, and someone opens it. How will you make sure they can find the other chunks of the document if they want them? Or at least know what original document they came from? How will they be able to tell it's an EPA document? [9]

The goal is to make sure anyone that reaches a section of a document knows what document it belongs to and can get the rest of the document if it's wanted.

What's the best way to present document source information?

There are three ways to present this information:

- Attach a cover sheet. You can create a template for your program. See Appendix B for a sample cover sheet.
- Add this information to the top of the first page of each chunk.
- Add this information to the document as a header and/or footer.

You can add this information in the original program (e.g., Word) or you can add it in Adobe Acrobat to a document that was produced in any program.

Document > Add Headers & Footers

---

[9] It may be possible to hide individual chunks from search engines.

Split up the document into chunks and place them into a separate directory on the server that contains a .noindex tag. The .noindex tag is used to build a robots.txt file. That file is used by search engines such as Google, etc. The robots file tells them which directories they should not index. So any directory with a .noindex tag is not indexed. This method makes the PDF chunks invisible and the search only finds the whole PDF or the HTML page. The HTML page will contain links to all the chunks.

Unfortunately, not every search engine will respect the robots.txt file.

No matter how your customer obtains a PDF chunk, they should be able to tell where it came from.

Lets you add formatted text in three columns. You can add a header or footer or both. You can make the information appear on all pages or on just one page.

Document > Add Watermark & Background

Lets you add formatted text or an image (jpg, bmp, or PDF) to a document. You can position this info in the margins and have it appear on all pages or on just one page.

In addition, be sure to fill in the Document Properties/Metadata fields for each file. For a chunk, you can use the same metadata as in the full document file, but add the specific information for that chunk (e.g., "Chapter 1, Calculation Methodology"). See Section 6 for more information on metadata.

Appendix G – PDF White Paper Timeline

Timeline (how we got here)

- **December 2002** - First time the PDF Workgroup appears on a Web Workgroup Meeting agenda
- **July 2003** - Introduced Guidelines and Best Practices at Web Workgroup Meeting in Seattle
- **December 2003** - Released Draft PDF Guidance and Short Checklist for Using PDFs on the Web
- **Early 2004** - Our linking best practices were added to the EPA Web guide, along with our other documents. We researched how to make PDFs accessible and concluded that sometimes it takes enormous effort and might not even be possible (using Adobe Acrobat 5 or 6, the versions available at the time).
- **Late 2004** - Jerry Carrillo, then EPA's Section 508 Coordinator, joined our workgroup, and helped us draft an undue burden justification.
- **December 2004** - We presented the draft undue burden justification to the Section 508 workgroup of the CIO Council. They convinced us that an undue burden justification was not appropriate, said that our proposed way of addressing PDF accessibility was 508 compliant, and encouraged us to develop a PDF policy for our agency.
- **Spring 2005** - EPA Web Workgroup members commented on the first draft of this white paper.
- **April 2005** - Adobe Acrobat accessibility expert Greg Pisocky gave his first presentations at an EPA Web Workgroup conference. He has presented hands-on training at every conference since then.
- **September 2005** - We presented our second draft of this white paper at the Web Workgroup conference and took further comments.
- **January 2006** - We tested different methods of linking to PDFs and tested them for accessibility and usability with blind and vision-impaired people.
- **April 2006** - We did follow-up testing of linking techniques with blind and vision-impaired people.
- **February 2010** – EPA's Web Council decided to stop chunking of PDF documents. A workgroup reviewed and updated the white paper to match current EPA standards. The document was updated for all PDF issues at the time.

PDF Subcommittee of the EPA Web Workgroup
Karen Reshkin, Chair
Amanda Babcock, Angela Palmer, Carol Cotton, Christine Dibble, Christine Tran, Donna Rickett, Jeff Tumarkin, Jennifer France, John Leitzke, Judy Dew, Karen Reshkin, Kathy Einwich, Lorien Wendt, Pat Burke, Peggy Heimbrock, Krista Carlson, Robert Johnson, Sally Martiny, Lawrence Lee

This subcommittee originated the EPA PDF white paper.

Updated by 2010 PDF Work Group of the EPA Web Council
Charles Bert, Judy Dew, Ross Gilleland, Michael Hessling, Jackie Kelly (BAH), Phoebe Stinson (BAH), Johnny Zimmerman-Ward (E² INC)

The 2010 Web Council PDF Workgroup updated the white paper to reflect current thought and practice at EPA but much of the original content and work of the PDF WWG Subcommittee is still reflected in this document.