

2006-05-05

Best practices for producing high quality PDF files

Formats Group, Deep Blue

<http://hdl.handle.net/2027.42/58005>



Best practices for producing high quality PDF files

Version 2.1, 21 January 2009

Adobe's Portable Document Format (PDF) is a popular means of sharing documents, especially those formatted for print. Not coincidentally, PDF files—when prepared properly—also provide an excellent archival file format.

The two most common ways to create PDF files are with applications like Microsoft Word and by a scanning process that starts with printed materials. Below you'll find specific recommendations for producing high quality PDF files using each of these methods, using either Adobe Acrobat Standard or Professional.

General guidelines for creating high quality documents

Enhancing the usability and appearance of your document via the following will do two important things. It will produce a more attractive print version and it will make your document easier to use for all users, including those with disabilities. These instructions assume you're using Microsoft Word; the same principles apply regardless of the program you use, though, and following these guidelines will also help you create more usable and accessible web pages if you're converting from Word to HTML.

- Use Word's "Styles" (found in the **Home** tab, or via the **Format** menu on the Mac) to identify headers, paragraphs, lists, and other structural elements. Create your own styles to add special formatting, rather than simply using the formatting toolbar to e.g. change fonts or italicize blocks of text.
- Include alternative text for all images and links, and place labels (such as "Figure 1.") near the images they refer to. You can do this in either Word or Acrobat when using Windows, and in Acrobat using a Macintosh.
- If you use a multi-columned layout, use the **Format > Columns...** feature in Word. Do not create 'virtual' columns using tabs, spaces, or tables.
- Create tables using either **Word's Insert > Tables > Table** (or, depending on the context, **Table > Insert Table** or **Table > Draw Table**) menu items to create a proper table. Do not "nest" data tables by putting one table inside another. Use a distinctive cell for each data entry.
- Place header and footer information (e.g., page numbers, notes, citations) in header and footer windows/areas rather than manually inserting them at the top or bottom of the page.
- For long documents, include a table of contents or other document guides with in-document links (bookmarks or anchors) that jump to specific sections of the text.
- For file names:
 - Use ASCII letters (a-z, A-Z), digits (0-9), underscores and hyphens ("_" and "-").
 - Do *not* use spaces, quotation marks, diacritic marks or other special/non-printing characters.
 - Reserve the period (".", full stop) for the file extension at the end of the filename.

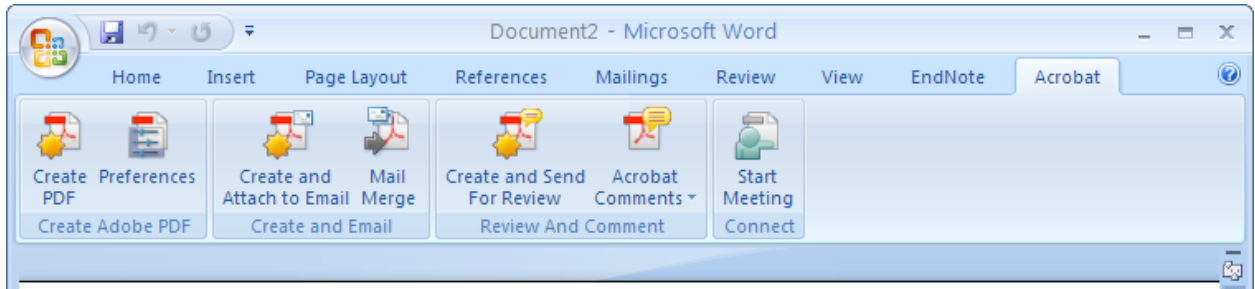
For more details on creating accessible and usable documents, consult the tutorials available at <http://wac.osu.edu>, from which portions of these recommendations have been adapted. You can also consult the Library's Knowledge Navigation Center <http://www.lib.umich.edu/knc/> for help setting up your documents. Adobe also provides detailed guidelines for usability accessibility at <http://www.adobe.com/accessibility/>.

Creating PDF files with MS Word and Adobe Acrobat on a Windows platform

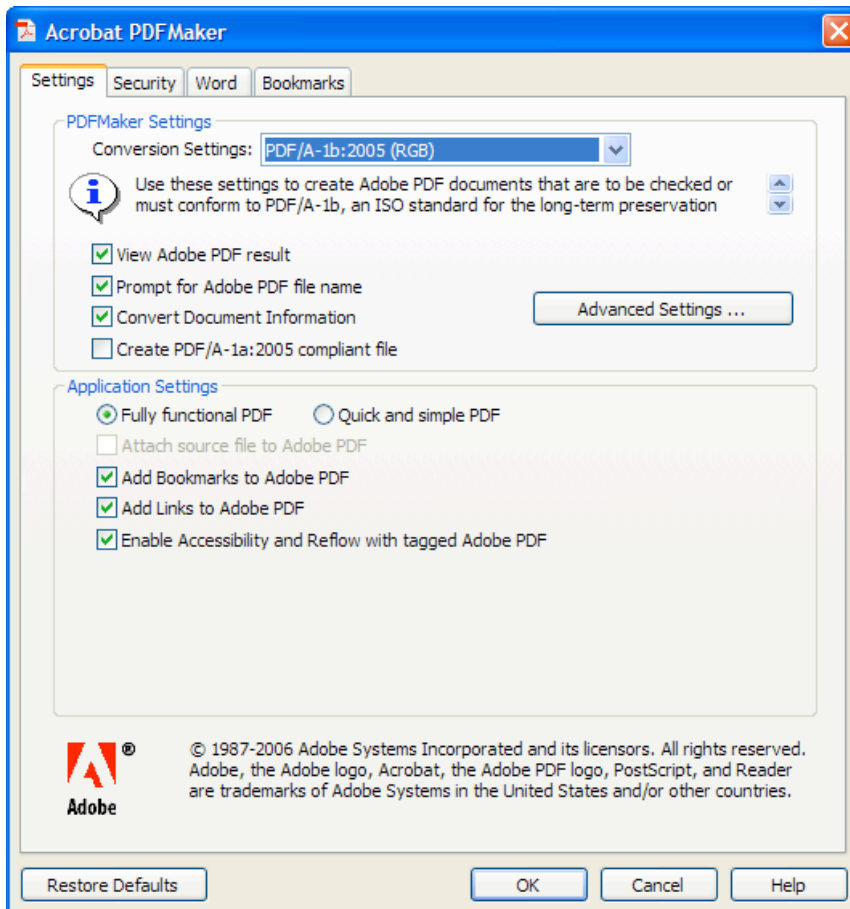
You can create PDF files from within Microsoft Word when you have the Adobe Acrobat Standard or Acrobat Professional software. *Note that Acrobat Standard and Acrobat Professional are not the same as Acrobat Reader, the latter being a free product with limited functions.* When configured properly, the resulting PDF files will be of high quality, suitable for deposit in Deep Blue.

Use the following steps to configure Adobe Acrobat 8 and create PDF files from within a supported application such as Microsoft Word 2007. (Steps 2-6 are very similar for earlier versions of Word and Acrobat.)

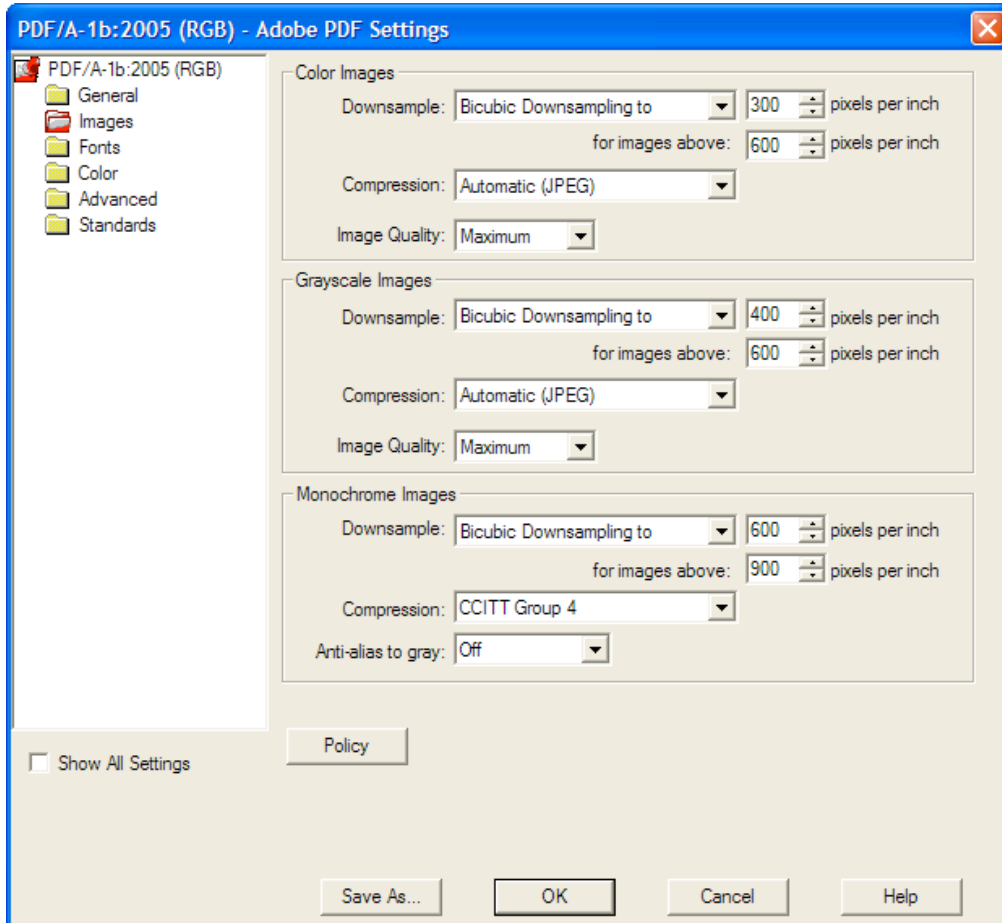
1. Under the **Acrobat** menu, select **Preferences** in the “**Create Adobe PDF**” section. A new window will appear, as shown below.



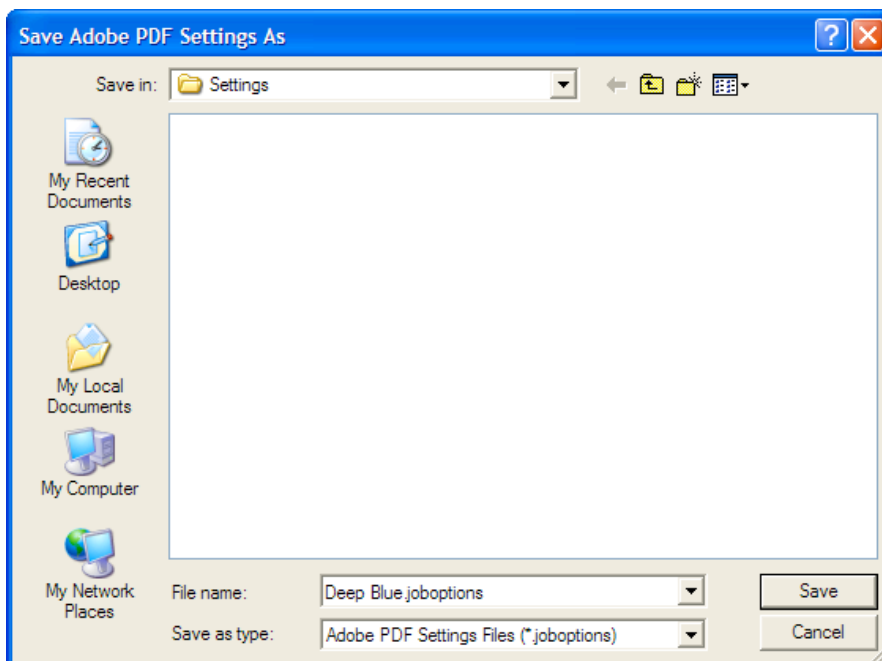
2. At **Conversion Settings**, select “PDF/A-1b:2005 (RGB)” and make sure you create a “Fully functional PDF” and that under that “Add Bookmarks to Adobe PDF”, “Add Links to Adobe PDF”, and “Enable Accessibility and Reflow with tagged Adobe PDF” are all selected. Then click **Advanced Settings**.



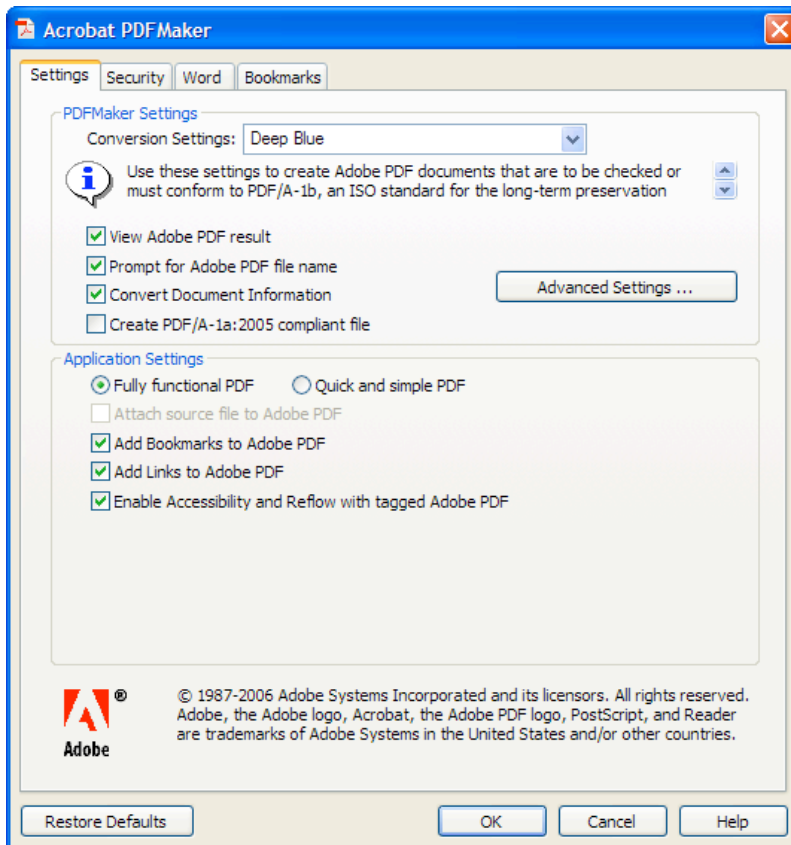
3. A new window will appear. In this window, click the **Images** folder and change all three of the pixels per inch settings, as shown.



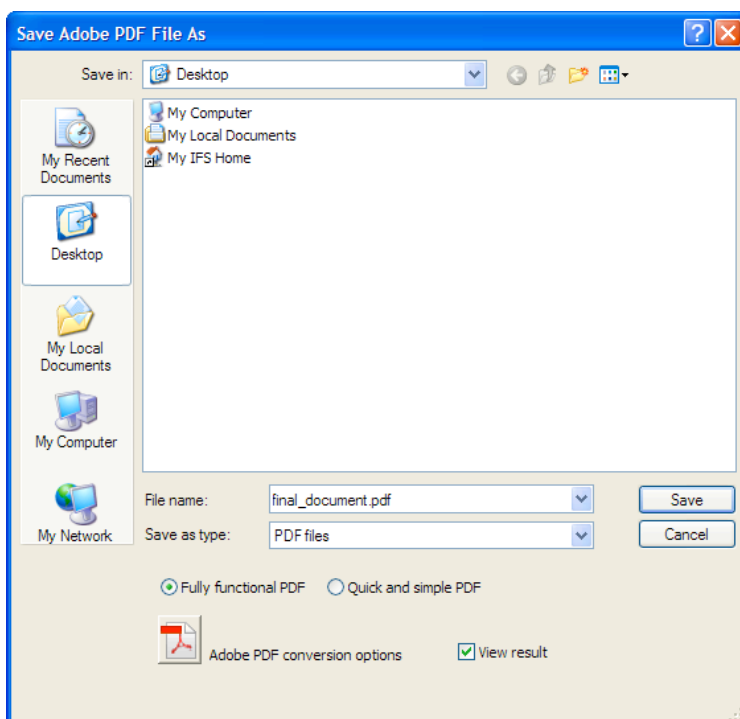
4. Click **OK**. You will be prompted to save the new settings. Enter “**Deep Blue.joboptions**” and click **Save** as shown.



5. Click **OK**. As shown, you will now have a special Deep Blue conversion setting available whenever you want to create PDF files for deposit.



6. To create a PDF from a finished document, under the **Acrobat** menu select **Create PDF**. You will be prompted to save the PDF file, as shown.



Creating PDF files with MS Word and Adobe Acrobat on a Macintosh

As per above, you can produce high quality PDFs using the Adobe Acrobat Standard or Acrobat Professional software. *Note that Acrobat Standard and Acrobat Professional are not the same as Acrobat Reader, the latter being a free product with limited functions.*

Microsoft Office 2008 and Acrobat Professional 8 (and 9) for the Macintosh do not currently work together to support automatic creation of PDF/A files. This is a known bug, anticipated to be addressed in a future version of Microsoft Office.

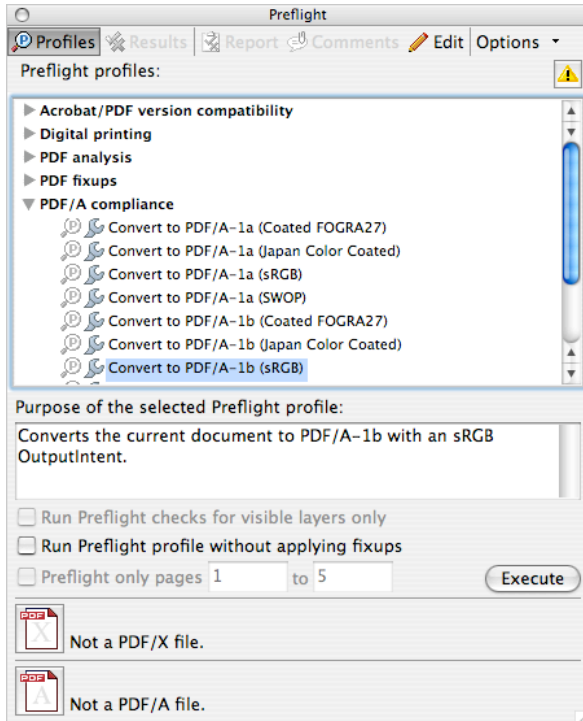
Fortunately, if you make sure your images are of the appropriate resolution then MS Word and Acrobat take care of *most* of the rest for you automatically. So, be sure to use the following settings for your images:

- Color and Grayscale: 400 pixels per inch minimum, JPEG compression w/Maxium image quality
- Monochrome: 600 pixels per inch, CCITT Group 4 compression or ZIP

With those in place, create the PDF by choosing **File > Save As...** and select "PDF".

The resulting file will *not* be fully compliant with the archival standard, but the essential features (embedded fonts, appropriate resolution for images) will be taken care of automatically. If you want to finish the process of creating a fully compliant PDF/A, you can do the following:

1. In Acrobat Professional, choose **Advanced > Preflight...**
2. Click on the arrow next to "PDF/A Compliance" and select "Convert to PDF/A-1b(sRGB)".



3. Click the "Execute" button and you will create a well-formed PDF/A, with "_A1b" appended to your file name, just before the .pdf extension.

Improve an existing PDF's usability and accessibility using Acrobat Professional

- Open your document in Acrobat Professional, and choose **File > Properties...** to determine whether the document has tags; those created using Microsoft Word on a Macintosh will not. If not, under the **Advanced menu**, choose **Accessibility > Add Tags to Document**. This will not solve all navigation problems, but it will significantly reduce them. (To check whether the document has tags before applying this, under the **View menu** select **Navigation Tabs > Tags**.)
- Provide alternative text for all images. You can do this using the "TouchUp Object" tool in Acrobat Professional. Select **Tools > Advanced Editing > Show Advanced Editing Toolbar**. Click on the far right icon. Click on the TouchUp Object tool at the far right. (As an alternative, just select **Tools > Advanced Editing > Touchup Object Tool**.)



Then, click on the image and then right-click (control-click for the Macintosh) and choose **Properties...** and then **Tag**. Fill in the "Alternate Text" box with descriptive information about the image.

- Make sure a language is specified for your document by choosing **File > Properties...** and then select the "Advanced" tab and set the Language field.
- Add bookmarks to further aid in navigation. These don't fully substitute for tags, or for structuring the document so that there's a clear, logical reading order. But they will assist all readers in getting to the information they want and need faster.
- To complete the process, under the **Advanced menu**, choose **Accessibility > Full Check...** and correct/fix any further accessibility problems identified. The report generated during the check will help guide you through those corrections or changes you might need to make.
- If you're using security settings to restrict copying, make sure that the "Enable Text Access for Screen Reader Devices for the Visually Impaired" option is checked.

To create PDF files by scanning print material: General principles

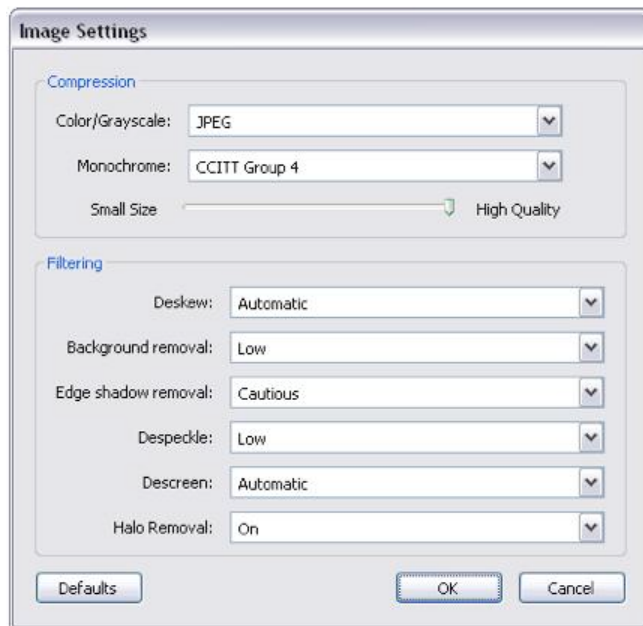
PDF files may also be created by scanning print materials. From almost every perspective, such methods are less preferable than creating PDF files directly from an application as described above due to increased file size and compromised viewing, printing, and usability characteristics. Nevertheless, PDF files created from scanning are a reliable way to preserve materials when no other options are available. Regardless of the method, the produced PDF file should have the following properties:

- All materials should be scanned at 100% scale to the dimensions of the original.
- Pages containing text and/or line art should be monochrome (black and white), 600dpi (or 300dpi for oversized materials greater than 11x16 in/28x41 cm), and be compressed using ITU Group IV compression.
- Pages containing photographs and/or illustrations should be 24-bit color using the sRGB color space or 8-bit grayscale, 400dpi (or 300dpi for oversized materials greater than 11x16 in/28x41 cm), and be compressed with JPEG compression using the highest quality setting.
- Missing or blank pages should be represented as blank images of the same size as the original.
- The PDF file should be optimized and should be in ("Searchable Image (Exact)" / "Image+Text") format.
- If your process creates intermediate files (TIFF, PostScript, etc), it must not change the resolution of the original scans (downsampling) or use lossy compression such as JPEG. (These methods, although they do reduce file size, will cause irreversible quality loss.) TIFF is an excellent choice for an intermediate file format, and acceptable compression schemes (which can all be used in combination with TIFF) are ITU Group 4 (for black and white material only), LZW, and Flate (Zip).
- If possible, embed OCR (optical character recognition) generated text in the document.
- The PDF should have tags and bookmarks as described elsewhere in this document.

To create PDF files by scanning print material: Using Adobe Acrobat Professional

Adobe Acrobat Professional has built-in scanning capabilities that can be used to create PDF files using printed materials. *Note that Acrobat Professional is not the same as Acrobat Reader or Acrobat Standard; only Acrobat Professional has support for scanning.* Use the following steps to create PDF files using a scanner with Adobe Acrobat Professional. These steps apply to both Windows and Macintosh, with only slight variations in appearance:

1. From the **File** menu, select **Create PDF** and then **From Scanner**. Select your scanner device, choose **Front Sides** or **Both Sides** as appropriate, select **Recognize Text Using OCR**, and select **Add Tags To Document**.
2. Click **Image Settings**. A new window will appear. In this window, change the settings as shown.



3. Click **Scan**.

Creating PDF files on a large scale

If you need to create a large number of PDF files, these approaches will not be efficient for you, although the principles outlined here apply equally to a large-volume operation. Before you start your project, contact us and we will be happy to provide advice to your staff or to your scanning vendor to make sure that your materials are in the best possible form for preservation. The Library's Knowledge Navigation Center <<http://www.lib.umich.edu/knc/>> can help you get started with this.

Questions?

If you have any questions, please contact us at deepblue@umich.edu and we will be happy to help you create high quality PDF files.

