

Sharing PDF Files Online

Donna L. Baker, ACE

*There's been a lot of Internet kerfuffle over the years regarding how to use PDF documents--or even whether to use them at all! Web usability pioneer Jakob Nielsen **has long promoted** limiting the use of PDF files online to those items users are likely to print. Although his recommendations don't consider the unique qualities of PDF, there are valid usability issues everyone should consider when moving PDF content online.*

Several years ago, the Internet community was abuzz with discussions of how to use PDF files online. Spearheading the naysayers, Dr. Jakob Nielsen suggested making PDF content available only when a user needs to print material, and not to use the format for online reading. That was 2003.

In Nielsen's **2007 update** to his "Top Ten Mistakes in Web Design" article, PDF files for online reading rank Number 2, second only to bad search tools.

In spite of the growth of PDF in this decade, and the increase in user education opportunities such as Acrobatusers.com, Dr. Nielsen still maintains PDF is a big usability problem.

I agree that using PDF online can create information flow issues. What Nielsen has failed to understand or promote, in my humble opinion, is how authors must handle PDF files for a simple, virtually invisible user experience.

So what IS the issue?

The PDF usability issue arose when users clicked a link that launched a PDF file and a viewer like Adobe or Acrobat Reader without their knowledge. Not only was a helper application required, downloading the file for viewing tied up the browser, thus halting the user's experience. Further, the layout displays at a size optimized for printing, not online viewing. Worst of all, according to Nielsen, "PDF is an undifferentiated blob of content that's hard to navigate."

His solutions included reserving PDF for distributing large documents intended for printing only. To read PDF content online, the document needed converting to a "real" Web page, that is, into HTML.

There's more, but I think you get the gist of the issue. Yes, PDF is different from an HTML page. Yes, there are situations in which it's best to use one format or another. But there are certainly ways to make the two formats behave and get along.



Start with the right conversion options

I'm sure you've all experienced this issue: You're browsing online and want to check out details of a company's product. You click the link to open the PDF file and then wait. And wait. Wait long enough, and you see the document. If not--a potential sale may be lost.

When you're converting source files to PDF, and know you'll be using the documents online, don't use conversion settings intended for high quality printing or professional print shop output. It's a waste of resources, including your prospective reader's time. And, if you produce a full-color brochure sent as a PDF file to your printer for output, never use the same PDF file online.

You don't always know where you'll distribute PDF documents when you first produce them. On the other hand, if you're splitting a catalog into product lines to download from your company Web site, use the *Smallest File Size* PDFMaker conversion settings designed for quick downloads.

Check your files

Regardless of how you intend to distribute a PDF document, take a minute and clean it up. In order of the amount of time and thought required, here are your options:

- **Save As:** Choose **File > Save As** and resave the file using the same name. Acrobat removes automatically saved file iterations and cleans up unused objects.
- **Reduce File Size:** Choose **Document > Reduce File Size** to display the Reduce File Size dialog box. Choose a compatibility version, or leave the default Acrobat 7.0 and later option. The newer the version, the greater the possible savings. Of course, you have to balance the savings against the probable program versions your audience is likely to use.
- **Optimize the PDF:** Choose **Advanced > PDF Optimizer** to open the dialog box. Don't be intimidated by the myriad of choices. You're trying to streamline your document. Click the Audit space usage button to see what elements in the document contribute to the overall size and whether you can make any savings (Figure 1).

FIGURE 1: Evaluate a document's contents for potential file size savings.

Optimize the contents

Figure 1 shows four areas of potential savings, including images, fonts, link annotations, and embedded files. The example file has internal and external links, and there's not much opportunity to remove those without disrupting the document's function. The same applies to the embedded files. There are three attachments, and they'll stay attached.

On the other hand, images make up over 20 percent of the document total, and fonts make up just less than 20 percent. Both of those items can (and should) be adjusted:

Description	Bytes	Percentage
Images	331,952	20.30 %
Bookmarks	998	0.06 %
Content Streams	2,755	0.17 %
Fonts	324,024	19.82 %
Structure Info	12,537	0.77 %
Acrobat Forms	91	0.01 %
Link Annotations	519,734	31.79 %
Comments and 3D Content	766	0.05 %
Document Overhead	6,669	0.41 %
Color Spaces	2,945	0.18 %
X Object Forms	3,482	0.21 %
Extended Graphics States	181	0.01 %
Piece Information	146	0.01 %
Embedded Files	428,829	26.23 %
Total	1,635,109	100 %

OK

- On the Images tab, type a downsampling value for the types of images used in the document. You don't need a 300dpi image for viewing onscreen when screen resolutions display 96dpi or less. You can downsample color, grayscale, and monochrome images separately.
- On the Fonts tab, unembed all fonts. An embedded font adds needlessly to file size. While any available font displays online in a PDF page (unlike HTML), unless you're bound by corporate policy to use a specific font, use a common font that Acrobat can display natively.
- On the Discard Objects tab, select items that won't interfere with the function of the PDF. For example, you can safely discard embedded page thumbnails or detect and merge image fragments without affecting the document's function.
- On the Discard User Data tab, select items that aren't necessary for

distribution, such as private data of other applications and hidden layer content.

- Finally, on the Clean Up tab, select all the choices. As you see in Figure 2, the clean up removes extraneous material, and optimizes the page content and the overall file for fast Web viewing (Figure 2).

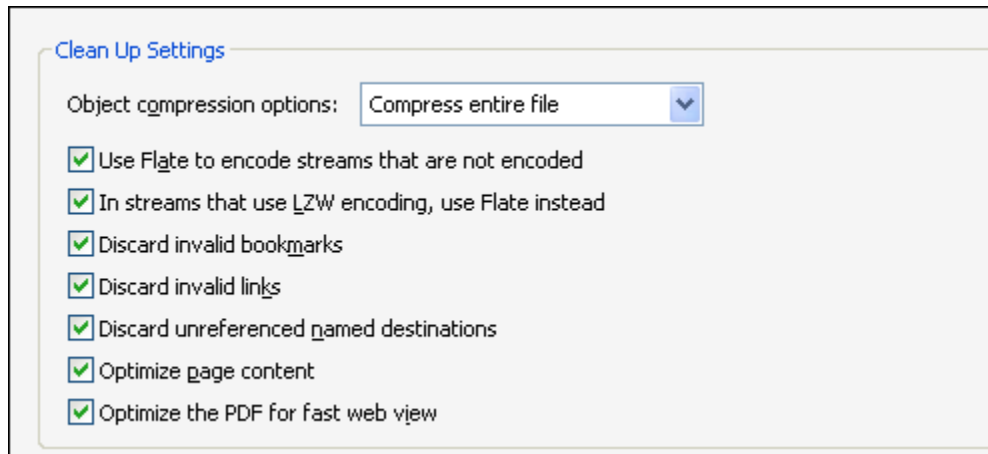


FIGURE 2: Clean up a range of unnecessary material.

After optimizing, check the results again in the PDF Optimizer. In the example file, while the image load didn't change appreciably, unembedding fonts saved about five percent of the file size.

Note: *The Optimize For Fast Web View is one of the most important settings to use for preparing online content. Acrobat restructures the document for byte-serving, a process where the file downloads page by page for quick display.*

Provide information for users and search engines

With the file optimized, it's time to modify some properties for effective online use. How many times have you downloaded an information manual of some type only to find there's no bookmark structure available? It's frustrating, time-consuming, and unnecessary.

Lack of navigation is another bone of contention cited by the anti-PDF folk. Again, it's a valid issue, but one with a simple remedy. Sure, it takes time to add and configure bookmarks, but if providing information to readers is the point of the exercise, take the time to offer readers a means to locate content easily. It's unnecessary to bookmark each topic or heading level, but at a minimum, add bookmarks for each section title (Figure 3).

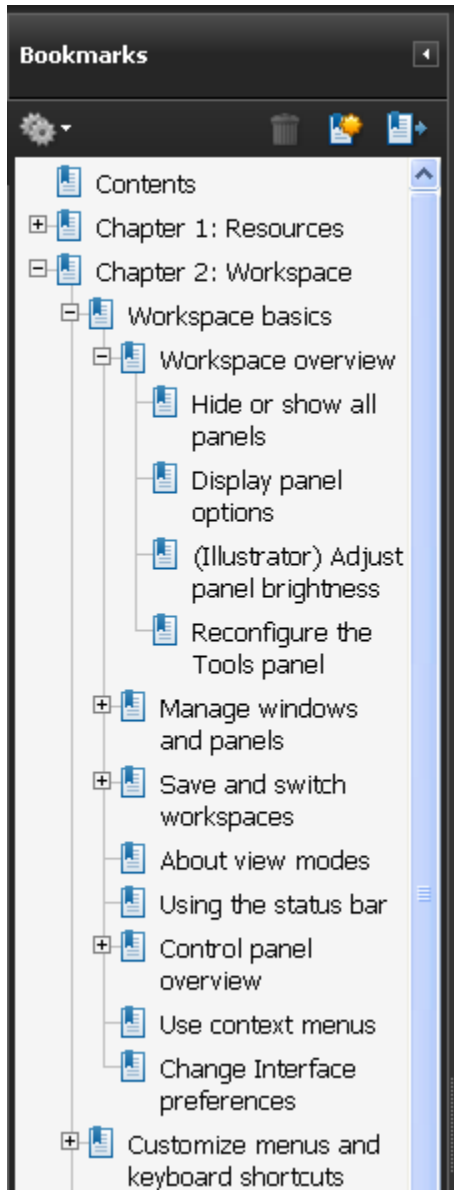


FIGURE 3: Insert the appropriate bookmark hierarchy.

Don't over-bookmark

The depth of bookmark levels corresponds with the information retrieval needs of the reader. Figure 3 shows a portion of the bookmarks for the InDesign CS4 Help Files. Users are likely to drill down to the third or fourth level bookmarks in this document looking for a particular topic. On the other hand, bookmarks for a line of upholstered furniture may need only two levels, with model numbers as child bookmarks of a parent product type bookmark.

With a navigation structure in place, head into the Properties dialog box to make some final adjustments. Choose **File > Properties** to open the dialog box. Here are my recommendations:

- On the Description tab, be sure to add a title, author and subject. Include some keywords for search engines.
- On the Security tab, add a password to protect your files. Interestingly, one of the comments in a Nielsen study of usability was that users didn't like PDF files since they couldn't readily copy and paste content from the page into their own work. Personally, I'd like to keep it that way.

On the Initial View tab, choose to display the Bookmarks panel and Page in the Navigation tab options. That way, when your document opens your readers see the navigation structure as well as the document.

Checklist for online file use

Before posting PDF documents online, be sure to:

- Clean up the files
- Unembed fonts
- Downsample images
- Add bookmarks and set initial view
- Insert document information for search engines
- Secure the document