

TITLE - Metadata: Hidden Information in Your Documents
AUTHOR - Alicia A. Slade
PUBLICATION - PA Family Lawyer
ISSUE/DATE – September 2005

If you have ever emailed a client or opposing counsel a Microsoft Word or Corel WordPerfect document, then you may have also sent confidential information along with the document. As you create and edit documents, hidden tags are embedded within the file that track information about you and the edits you have made. This hidden information is called Metadata, which is “data about data.” Metadata can cause not only embarrassing situations, but also reveal potentially damaging information.

Microsoft stores metadata in Word, Excel, and PowerPoint files. The following is a partial list of the types of hidden data stored:

- the name and initials of the person who created the file;
- the firm or organization name;
- the name of your pc;
- personal summary information;
- comments;
- deleted text;
- revision marks;
- previous authors and editors;
- document versions;
- hidden text;
- the hard drive/location where the file is saved;
- printer information;
- macros and templates.

Corel also stores similar information within WordPerfect files. Most software programs track some form of metadata within the files. Therefore, if you share documents via email, diskette, CD, or in any electronic form, then the files can contain metadata.

There are several approaches that can be taken in order to minimize metadata within your documents. First, you could attempt to manually remove the metadata. This can be extremely cumbersome and some of the steps will be discussed below. Next, you could scan your files to a PDF file or convert the file to a PDF format. You must be careful because PDF files also contain a minimal amount of metadata. Also, if your file has been converted to a PDF file and not scanned, then the recipient could convert the file back to Word and potentially see the original metadata in the file. Another option is to download a free add-in from Microsoft called Remove Hidden Data which will add a utility to clean your documents of metadata from the menu. Unfortunately, there is no such tool for WordPerfect at this time. Lastly, you could purchase a “data scrubber” software product which will “clean” your documents before emailing them.

There are several steps that you can perform to manually change metadata information within Word. You can change your company information from appearing on documents by modifying your user information. Select Tools, Options, User Information. To not include personal information when you save a document, select Tools, Options. From the security tab, select the “remove personal information from file properties on save” check box. Consider disabling “fast saves” by select Tools, Options, Save. Deselect the “fast saves” checkbox. If you are tracking changes, be sure to remove all redlining information from the document prior to saving. You can also remove old versions of the document by selecting File, Versions. Select and delete different versions of the document. To open the Properties information about the file, click File, Properties. On the Summary tab, you will see the title, author, company information. The General tab will show the location of where the document is stored and the creation date/time, modified date/time and accessed date/time.

Microsoft offers a Remove Hidden Data Tool that integrates with Microsoft Office XP and Office 2003. You can download the add-in and install it into your Microsoft Office. Visit www.microsoft.com/downloads and in the Search for a Download dialog box, enter the keywords, Remove Hidden Data, then click Search. The Office 2003/XP Add-in Remove Hidden Data will appear and you will be able to download this file. Once you have installed the Remove Hidden Data Add-in, open a document, then on the File Menu, click Remove Hidden Data. The tool will require you to save another copy of your document prior to it being cleaned. The following are examples of the type of metadata that will be removed: track changes, versioning, comments, hidden text, user names, and VB macros.

Metadata Scrubber, from BEC Legal Systems, offers an even more automated approach that works in conjunction with Microsoft Outlook and Microsoft Word 2000, XP, and 2003. When attaching a file to an email message, the scrubber will initiate the removal of the metadata upon sending the message. The Metadata Scrubber will remove track changes, versioning, comments, hidden text, user names, document properties, Visual Basic macros, etc. When running a Microsoft Exchange server, the Metadata scrubber can be customized to run in a silent mode with no user intervention required. The scrubber can also be run directly from Microsoft Word with the click of a button. This product offers you a way to remove hidden data which is much more reliable than removing it manually.

So, before clicking the Send button to send an email with an attachment, it is important to take the necessary precautions to verify the files that you are emailing do not contain sensitive metadata.