

Beware the Dangers of Metadata

Are you unwittingly sending confidential information to clients or opposing counsel? If you have e-mailed a Microsoft Word or Corel WordPerfect document to either, the answer to this question is likely yes.

When you create and edit your Word or WordPerfect documents, information about you and the edits you make is automatically created and hidden within the document file. This information is called metadata.

Litigators need be aware of metadata for an additional reason: It can provide extensive information for building an electronic paper trail as to who created, read, or even deleted text from a particular document, whose computer it was stored on, and more.

What is metadata?

Metadata can be simply described as "data about data." Think of it as a hidden level of extra information that is automatically created and embedded in a computer file.

On its Web site, Microsoft indicates that the following metadata may be stored in documents created in all versions of Word, Excel and PowerPoint:

- your name and initials (or those of the person that created the file);
- your firm or organization name;
- the name of your computer;
- the name of the local hard drive or network server where you saved the document;
- the name and type of the printer you printed the document on;
- other file properties and summary information (see below);
- non-visible portions of embedded OLE objects;

- the names of previous document authors;
- document revisions, including deleted text that is no longer visible on the screen;
- document versions;
- information about any template used to create the file;
- hidden text; and
- comments.

Similar metadata exists within WordPerfect files, and metadata data security issues affect most software programs.

While some metadata can easily be viewed within the program that has created a file, in most circumstances hidden metadata can usually only be seen by using a binary file editor. However, sometimes hidden metadata can become visible accidentally – for example, when WordPerfect opens and improperly converts a Word file, or when a corrupted file is opened. In these instances, both of which are quite possible in a law office, both the normally visible text and hidden metadata can appear on a computer screen.

Metadata is very useful in some circumstances. Document management and precedent systems make extensive use of metadata. Information such as: who created or edited a document; the document's usage and distribution history; and relevant keywords, subject or matter information, is essential for helping other users find a relevant document within a precedent databank or other document repository.

The problem with metadata, especially for lawyers, comes when people electronically share files as an attachment via e-mail, on a floppy disk or CD-ROM, over a network, or through an extranet. Electronic

document files include both the information you see on the screen, and all the metadata you don't see. This metadata can often be confidential information, and can be potentially damaging or embarrassing if seen by the wrong eyes.

Metadata in Word

How can you view metadata in one of your Word documents? (WordPerfect users should jump ahead three paragraphs) Find and open a letter or agreement you recently emailed to a client or opposing counsel. Click on File, then Properties. This opens the Properties dialog box. It contains a variety of summary type information about the file.

On the General tab you can see what hard drive the document was stored on, and the time and date it was created, last modified and viewed. On the Summary tab you can see the name of the author, your firm name, as well as the name of the template that was used to create the document.

The Statistics tab contains information about the size and structure of the document, including the Total Editing Time in minutes. This statistic is really the total amount of time the file was open on a computer, regardless of whether someone was editing it or not. What if a client saw this information, and the time indicated was significantly less than the amount of time you docketed for working on this document? This discrepancy could be completely justifiable, but you could find yourself explaining it to an upset client.

Metadata in Wordperfect

In WordPerfect you can see the basic file summary type of metadata you see in Microsoft Office documents by selecting File, then Properties.

WordPerfect also has a feature called Undo/Redo History. It can allow you to view hundreds of past changes in terms of what text was cut, copied and even deleted from the document. Open a WordPerfect document. Click on Edit, then Undo/Redo. This opens the Undo/Redo History which lists past changes to that document, assuming Undo/Redo is turned on. Click on the Options button, and then uncheck "Save Undo/Redo items with document" to turn it off. Look at some of your WordPerfect files to see if you can view summary metadata or the Undo/Redo History.

In many instances lawyers will reuse and adapt a document they created for a previous client. This makes perfect sense from an efficiency point of view. However, note that deleted text can remain within a document. This is especially a danger if you use features that track document changes, allow for commenting, or keep versions of documents. What would happen if your client sees confidential information about the client the document was originally created for, or if opposing counsel saw changes that were made in an agreement at the drafting or client review stage?

How do you remove metadata?

Being aware of metadata is just the start. Obviously you must take steps to reduce or eliminate the metadata in your documents. Sending a fax or paper copy by regular mail would solve the problem, but will likely not be an option in many circumstances. If you want your client to review and edit a document, sending it electronically is the only practical option. Indeed, in many cases, clients, opposing counsel and even the courts expect to receive documents electronically.

You have a number of options to reduce or eliminate metadata from your documents.

Word, PowerPoint and Excel users should turn off the Fast Saves feature. To do this click on Tools, then Options, then the Save tab, and uncheck "Allow fast saves." In older versions of Microsoft Office products it will be turned on by default. This feature lets a computer more quickly save a file by not removing deleted text from it. When computers were much

slower it was perhaps a helpful feature. With the more powerful computers we now have, you won't notice any difference with this feature turned off.

If you used features such as tracked changes, document versions or comments, make sure you delete the information that is being kept within the document with these features.

Office XP includes some new features to help reduce the accidental disclosure of metadata. Even more features are included in Word 2003 and the other Office 2003 applications. They now include a Security tab in the Options dialog box (select Tools, then Options to view this tab). You can specify that some metadata not be saved in a document in this dialog box. The Information Rights Management feature in Office 2003 can also be used to reduce paper trail types of metadata being stored within documents.

Converting files to PDF format with Adobe Acrobat or other PDF creators will strip out most metadata. For this reason many have adopted a practice of sending only PDF documents to clients or opposing counsel, especially if the recipient doesn't need to edit the document. Sending documents in a format that prevents a document from being changed can actually be helpful or even necessary in some cases. Ideally you should clean out metadata such as tracked changes, versions and comments as it can end up being included in the PDF file after conversion.

While converting a file to PDF format will help strip out metadata from the original document, remember that PDF files can also contain their own metadata. This is usually basic information such as the name of the person who created the file, date of creation, file location etc. Select File, then Document Properties to view the summary metadata information within a PDF file. In this same dialog box you can add further restrictions on how the document can be accessed, used, copied and printed in the Security Options settings.

If you want the recipient to edit the document, you need to send it in its native format, but without metadata.

There are several programs that can help

you identify and clean metadata from your documents.

In February this year Microsoft released the Remove Hidden Data Add-In. It will permanently remove hidden and collaboration data, such as change tracking and comments, from Word, Excel, and PowerPoint files in Office XP and Office 2003 only.

For Word, Excel and PowerPoint documents, the most widely used metadata scrubber is the Metadata Assistant, sold by Payne Consulting Group (www.payneconsulting.com) for US\$79. A free demo version of the program will show you the metadata within a Word document, but won't clean it out.

Other metadata removal programs for the Microsoft suite of products include Out-of-Sight by SoftWise Consulting (www.softwise.net), ezClean from KKL Software (www.kkl.com) and Workshare Protect by Workshare (www.workshare.net). For Word only there is Doc Scrubber (www.docscrubber.com).

Unfortunately there is no software program for easily and automatically removing metadata from WordPerfect documents.

For more information on metadata, see the following resources:

- Word, Excel and PowerPoint users should visit the Microsoft support page at <http://support.microsoft.com/>. For more detailed information on removing metadata from Word 97, 2000 or 2002, see respectively, Knowledge Base articles 223790, 237361 or 290945.
- WordPerfect users should visit the Corel knowledge base at <http://support.corel.com/> and search for "minimizing metadata."

Take some time to understand more about metadata, and how you can take steps to eliminate it from your documents. This will help you prevent an unwanted or embarrassing disclosure of confidential information.

Dan Pinnington is Director of practicePRO. You can reach Dan at dan.pinnington@lawpro.ca.