

Metadata: The Hidden Disaster That's Right in Front of You

What Is Metadata?

Succinctly defined, metadata is “data about data.” Metadata is embedded in all Microsoft Office documents

Microsoft Word, Excel and PowerPoint include automated features to aid in document production and collaboration. These features embed electronic information (metadata) in a file, which can reveal the identity of those who edited the document (revision authors); track the time, date, and frequency of edits (track changes and revisions); reveal inserted comments and the document template; and other data employed to control the document's text and format. Metadata is placed in a document by the operating system, the application, and by users utilizing the automated features of the application.

The metadata contained in a Word document doesn't necessarily create risk of adverse disclosure. In fact some document metadata is necessary for formatting or automation macros within a document. Some document metadata, such as tracked changes, may be used to collaborate with co-counsel, but one might not wish to share such information with one's adversary. The commonly held opinion is that information should be removed before a file is shared outside a firm's electronic walls to avoid violating attorney-client privilege, disclosing sensitive information to third parties and so on.

Before determining how your law office is going to manage metadata, it is important to understand the basic facts about document metadata.

Fact 1: Metadata Exists in ALL Microsoft Office Documents

A rule of thumb when considering metadata is that every time a document is opened, edited and saved, metadata is added by the operating system, the application itself, and through the use of certain automation features.

Some firms claim that they do not have a “metadata problem” when in fact ALL Microsoft Office documents contain some kind of metadata. The question is whether the metadata revealed is harmful or not. It is always better to err on the side of caution.

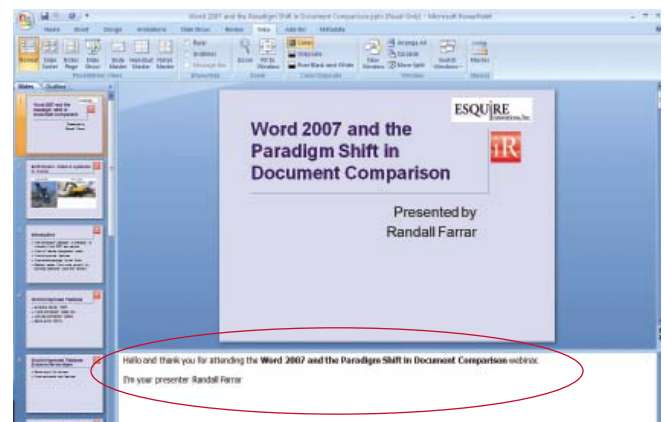
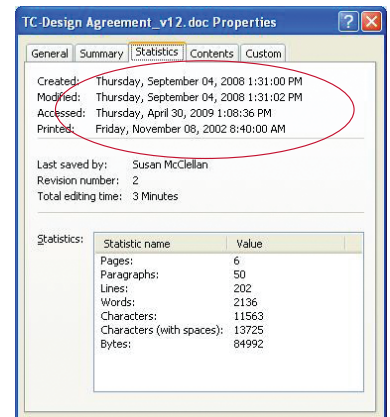
Fact 2: Metadata Can Be Useful

Microsoft Word metadata is often essential to the document production process to automate formatting and reduce editing and collaboration time. For example, the date fields (under document properties) are referenced when searching for documents created in a specified time frame, or

to gain quick access to documents from “My Recent Documents.”

Tracked changes can be useful when editing a document with multiple co-counsel or colleagues to identify which editors have made specific changes. In Excel, metadata can also be very useful and includes formulas in a spreadsheet, hidden columns, author names and creation dates of documents. In PowerPoint, metadata includes author information and presentation creation dates, as well as speaker notes and links to graphs or other statistics from outside documents.

RANDALL FARRAR (randall.farrar@esqjnc.com) is the president of Esquire Innovations, Inc., a software company that develops Microsoft Office integrated applications for the legal market, located in Temecula, CA.



Fact 3: Metadata Can Be Harmful

Metadata can be harmful when users unknowingly send documents that contain confidential or potentially embarrassing information. There have been many well-publicized cases in which tracked changes or hidden comments have been left in a document sent via email or shared on the Internet. Two examples of high profile metadata blunders are the SCO Group's lawsuit against DaimlerChrysler and a United Nations report.

A Microsoft Word document from SCO's suit against DaimlerChrysler originally identified Bank of America as the defendant instead of the automaker. Metadata revealed that SCO spent considerable time building a case against the bank before changing the name on the suit to

DaimlerChrysler. More information can be found at the following web link: http://news.cnet.com/2100-7344_3-5170073.html.

In a United Nations report, tracked changes were discovered in a document that supported the published conclusion that Syria was behind an assassination in Beirut. Confidential and sensitive information as well as evidence that the report may have been altered after it was submitted to the United Nations were disclosed. More information can be found at http://www.timesonline.co.uk/tol/news/world/middle_east/article581486.ece.

Law firms that deal with sensitive and confidential information on a daily basis must be diligent in managing their metadata or they too may find themselves the subject of media reports and embarrassment.

Fact 4: Tracked Changes Can Easily Be Left in a Document

Despite the far-reaching negative effects of metadata discovered in a document, something as simple as leaving tracked changes in a document can easily happen. Consider the following scenario.

An attorney switches on the "Track Changes" feature in Word to make edits to a document. After collaborating with his assistant and associates he is satisfied with the changes.



He decides to send it to the client for review and clicks on the "Review" ribbon in Word 2007 and changes the document to "Final" in the Tracking section.

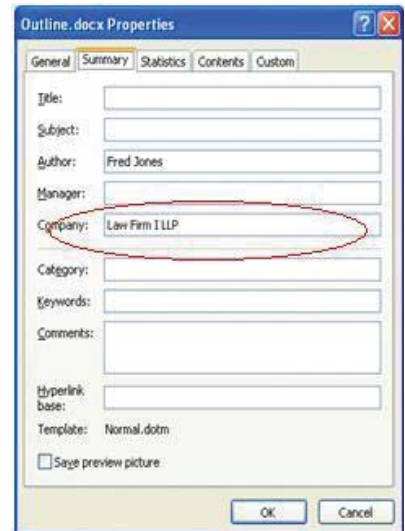
The tracked changes disappear from the document. He assumes they are no longer there, clicks on send via e-mail and forwards the document to his client. The client opens the document to see all of the tracked changes displayed. This occurred because the attorney did not accept all of the changes in the document; he merely hid them from view. When the client opened the document the "Display for Review" settings were set by default to "Final Showing Markup," thus revealing all of the changes in the document.

To make sure that this scenario does not occur, and that there are no tracked changes left in a document, always accept all changes.

Fact 5: Metadata Can Be Found in the Document Author Information

Multiple author names can remain with a document as it is edited and revised. Microsoft Word automatically pulls the author name from the User Information for the "Last saved by" author (found by accessing the Office Button then Word Options|Popular), and will save the names if there have been multiple editors of a document.

When a document is created from an earlier document using Save As, the author name from the original document will stay with the document as will the company name. Often an attorney will create new documents from legacy documents that could have been produced when working for a previous firm. Unless the company information is manually updated by the user, or cleaned by a metadata software application, it will stay with the document.



If a law firm regularly uses the same document for multiple clients and/or uses documents created by lawyers when they were employed by previous firms, the client could see a different author, law firm and client listed in the properties. This information could lead to serious questions from a client as to a firm's billing practices. However, there are ways to control author information on docu-

ments. Microsoft Word has five areas that collect author information:

- User Name
- User Initials
- Document Author
- Manager
- Last Author

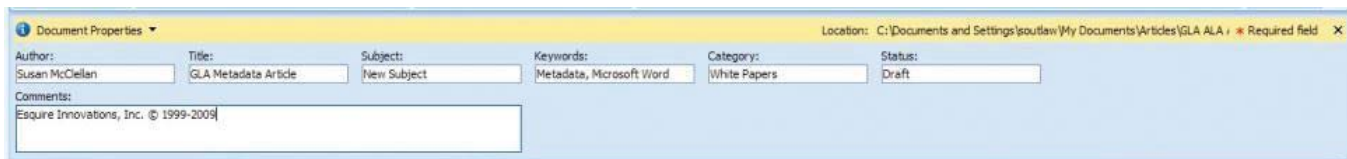
The User Name and User Initials control what appears in the author properties of a Microsoft Word document. User Name and User Initials are found in Word Option|Popular|Personalize, depending on your copy of Microsoft Office.

Microsoft Word documents also contain other properties that reveal the document author, which can be found in the built-in document properties of a document.

To view these properties click on the Office button select Prepare|Properties. A display bar will open at the top of the document.

The document author is pulled from the "Word Options" settings described above and inserted when the document is created. This stays with the document until it is changed or deleted.

The other fields displayed are user input properties. That means one has to manually place text here. Some



template and macro applications use this field for automation purposes and place information in these properties. Unless the firm is using an automated metadata software, be aware of these properties and that they will remain with the document until they are changed or deleted.

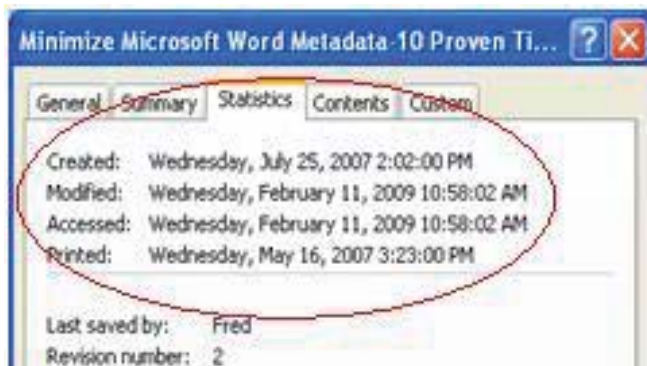
Fact 6: Metadata Is a Document's Dates and Times

In the Microsoft Word "Statistics" tab the Created, Modified, Accessed and Printed fields are displayed. This information can cause potential problems for a law firm.

For example, an attorney is creating a new contract for a client. The contract requires some standard language. The attorney has prepared similar contracts before, so she opens up a contract that she had created in Microsoft Word for another client when she worked at a different firm. The attorney makes edits as needed and e-mails the contract to her client. Upon receipt, the client opens the document and, since she has heard about metadata, opens "File Properties" to view any data. File properties can be accessed in Office 2007 by clicking on Office Button | Prepare | Document Properties | Advanced Properties. By viewing the Statistics tab the client sees a Created date of Wednesday, July 25, 2007, one year before she was a client and a Modified date of Wednesday, February 11, 2009, which is the current date.

Even more puzzling is the Printed date, which is several years earlier, indicating that the last time this document was printed was Wednesday, May 16, 2007. This date will remain unchanged until the document is printed again.

Word files can contain a history that reveals the true age of a document. That history will stay with the document until it is "cleaned" using a metadata management tool.



Metadata of this type can be useful when searching for documents created in a specified time frame, or to gain quick access to documents from, for example, My Recent Documents. But a firm may not wish to reveal this type of information to a client being billed an hourly rate for creating the document.

Fact 7: There Are More Than 200 Types of Document Metadata

There can be more than 200 types of metadata added to a document.

In addition to the examples cited above, less commonly known metadata include:

Field Codes – Naming conventions for custom field codes may disclose information about the drafting process not disclosed by the text.

Bookmarks – Naming conventions for bookmarks may disclose information about the drafting process.

Routing Slips – When the File | Send | Routing Recipient function is used, the recipients' email addresses are stored in Word's electronic file (not available in Office 2007).

Firm Styles – Custom style names can sometimes be firm specific and therefore considered metadata.

Prevent Metadata Issues – Establish a Metadata Policy

Law firms, more than most users of Microsoft products, can be embarrassed – or worse – if metadata is not properly managed. Each law firm should have a metadata policy that is utilized by all attorneys and staff who work on firm documents. Considerations to take into account when establishing a metadata policy include:

- Educate yourself and your users about metadata.
- Review the applicable New York opinions (and those of other states and entities, as needed) regarding metadata.
- Review firm documents (on internal networks and published on external networks). Is your firm inadvertently sharing confidential information?
- Involve attorneys and your IT department and establish a firm approach based on your findings.
- If necessary, bring in a consultant to advise your firm on a metadata policy.
- Periodically review the firm's policy to address any new rulings on metadata and/or changes to Microsoft.

Enforcing the Policy

All firms should consider purchasing metadata management software. The software should be flexible enough to execute firm policy, automated enough to enforce firm policy and easy enough for users to understand and utilize.

The latest Microsoft Office program includes a metadata tool called Document Inspector. Since Microsoft applications add metadata to files, it presents a somewhat contradictory position for Microsoft to provide a tool for removing that metadata. Firms who already practice a metadata policy have found that the main weakness with

Document Inspector is the lack of automation. The onus is on individual users to “inspect” documents and then decide which metadata to remove. This approach proves ineffective in enforcing a metadata policy throughout an organization. Metadata management software, on the other hand, removes metadata more thoroughly and is designed to help firms automate and therefore enforce metadata policies. The most popular products available for metadata management can be found by searching for “metadata management software” in Google.

The success of any policy hinges on the execution. A firm’s metadata policy will be more successful if staff can grasp what metadata is, when it can be useful, when it can be harmful and how to manage the metadata in documents. Consider bringing in outside trainers to help educate your firm with hands-on training.

Metadata and New York Law Firms

Historically, opinions on whether there is a significant risk with metadata and if so what must be done to address that risk have varied among attorneys, IT departments, management, bar associations and other governing entities. In

the past few years, a multitude of governing bodies have drafted and issued opinions regarding metadata. New York has opinions specifically addressing an attorney’s ethical obligations regarding metadata in place. Law firms in New York should ensure they are in accordance.

Law associations throughout New York, including the New York State Bar Association, the New York City Bar Association and the New York County Lawyers’ Association, have released formal opinions on attorneys’ ethical responsibilities regarding metadata.

The New York State Bar Association’s Committee on Professional Ethics Opinion 749 and Opinion 782 state that a lawyer’s ethical obligations regarding metadata are summarized as follows:

Lawyers may not ethically use available technology to surreptitiously examine and trace e-mail and other electronic documents.¹

and

Lawyers must exercise reasonable care to prevent the disclosure of confidences and secrets contained in “metadata” in documents they transmit electronically to opposing counsel or other third parties.²

The New York State Bar Association has also developed a basic guide for attorneys regarding metadata, which outlines the legal and ethical issues for lawyers regarding metadata, how to preserve and produce metadata, and the ethical obligations specific to New York lawyers.³

The New York County Lawyers’ Association’s Professional Ethics Committee Opinion 738 states in part,

[A]ttorneys are advised to take due care in sending correspondence, contracts, or other documents electronically to opposing counsel by scrubbing the documents to ensure that they are free of metadata, such as tracked changes and other document property information.⁴

As more states sound off on metadata and an attorney’s responsibility, New York firms with practices in multiple states should also make sure that their policies are acceptable in every jurisdiction in which they practice. ■

1. New York State Bar Association, Committee on Professional Ethics: Opinion 749 (2001), available at http://www.nysba.org/AM/Template.cfm?Section=Ethics_Opinions&TEMPLATE=/CM/ContentDisplay.cfm&CONTENTID=6533.

2. New York State Bar Association, Committee on Professional Ethics: Opinion 782 (2001), available at http://www.nysba.org/AM/Template.cfm?Section=Ethics_Opinions&CONTENTID=6871&TEMPLATE=/CM/ContentDisplay.cfm.

3. “Metadata: Basic Guidance for New York Attorneys” was produced in April 2008 by the Committee on Electronic Discovery of the Commercial and Federal Litigation. The guide can be found at <http://www.nysba.org/Content/NavigationMenu4/Committees/Metadata.pdf>.

4. New York County Lawyers’ Association, Committee on Professional Ethics: Opinion 738 (2008), available at http://www.nycla.org/siteFiles/Publications/Publication1154_0.pdf.

Find us on the Web!

Just a click away:

The NYSBA *Journal* is available to you anytime at all.

Log in as a member:

Membership gives you access to current issues and the *Journal* archive on HeinOnline. The archive offers the *Journal* in a word-searchable format, beginning with the first issue in 1928.

Find an article:

Our word-searchable index lists all *Journal* articles from 2000 through present.

Let us know:

Comment on any article you’ve read, topics you’d like addressed or the issues facing today’s practitioners through the editor’s blog.

The *Journal* at www.nysba.org/barjournal.

The Editor’s blog at <http://nysbar.com/blogs/barjournal/>. Click on “comments.”