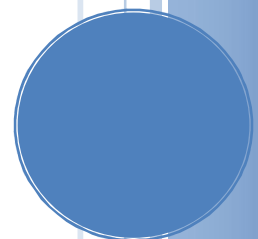


10 Facts about Document Metadata Part 1 of 2

By Randall Farrar
in the GLA ALA Leadership Exchange

Published
April/May 2010

ESQUIRE
INNOVATIONS, INC.



10 FACTS ABOUT DOCUMENT METADATA

PART 1 OF 2

What is Document Metadata?

Metadata is often defined as “data about data.” More specifically, the term metadata refers to data that is attached to a computer file beyond the readable content.

Microsoft Office (which includes Word, Excel, and PowerPoint) includes many automated features to aid in document production and collaboration. These features embed electronic information in a file that can reveal the identities 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.

There has been much written about the danger of inadvertently disclosing document metadata. Increasingly State Bar Associations have issued opinions regarding an attorney’s professional responsibility regarding metadata leading many law firms to address the management of metadata and develop specific policies or update existing ones to ensure they are in accordance.

What does the California State Bar Say?

THE STATE BAR OF CALIFORNIA, STANDING COMMITTEE ON PROFESSIONAL RESPONSIBILITY AND CONDUCT, FORMAL OPINION NO. 2007-174

“An attorney is ethically obligated, upon termination of employment, promptly to release to a client, at the client’s request” documents. Further “the attorney is ethically obligated to take reasonable steps to strip from each of these electronic items any metadata reflecting confidential information belonging to any other client.”

(Also see: THE STATE BAR OF CALIFORNIA STANDING COMMITTEE ON PROFESSIONAL RESPONSIBILITY AND CONDUCT FORMAL OPINION INTERIM NO. 05-0006)

The metadata contained in a Word document other than the intended text 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 would not want to be shared with adversarial counsel.

Before developing a metadata management policy the first step should be to understand what metadata is and how it is produced by common applications.

Here are Facts 1-5 to help you understand Document Metadata

1. METADATA EXISTS IN ALL MS OFFICE DOCUMENTS

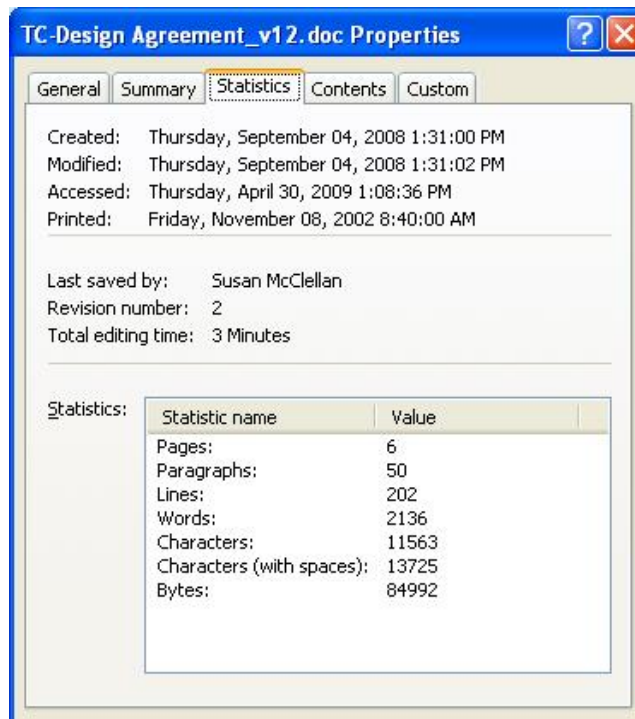
Microsoft Word, Excel and PowerPoint documents contain information (data) other than the actual content of the document.

This data is placed there by the operating system, the application itself and by utilizing certain automation features. Automation features aid in document production and collaboration, but can embed electronic data in a file including the identity of those who edited the document (revision authors); track the time, date, and frequency of edits (track changes and revisions); hidden comments, information about the document template and firm styles; and other data employed to control the document's text and format. All of this data is called "document metadata", and can be easily discovered if not removed.

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".



Date/Time Fields in Document Properties

Tracked changes can be useful when editing a document with multiple co-counsel or associates to identify which editors that have made specific changes.

THIS DESIGN AGREEMENT (the "Agreement") is entered into effective as of ~~November-December~~, 2005 (the "Effective Date"), by and between Worldwide Lint ~~LLC~~ *Inc.* ("Worldwide Lint"), a Texas ~~limited liability company~~ *corporation* limited liability company with an address at 5300 Town & Country Blvd., Suite 150, Frisco, TX 75033 and ~~ABCXYZ~~ Company ("ABCXYZ"), Vermont ~~cooperative~~ corporation with offices at 101 North Third Street, Moorhead, Vermont 56560.¶

RECITALS¶

~~A. Among other things, Worldwide Lint is currently developing a flash calcining procedure which rapidly turns spent Lint into marketable calcium oxide.¶~~

~~B.A.~~ → Worldwide Lint and ABCXYZ wish to explore the potential implementation of Worldwide Lint's flash calcining process at ABCXYZ's facility at Winchester, North Dakota (the "Winchester Facility").¶

~~C.B.~~ → The parties ~~recognize-understand~~ that certain design costs will be incurred by both parties prior to execution of the Definitive Agreements (as defined in Section 24 below) and wish to provide for the treatment thereof.¶

NOW, ~~THEREFORE~~, in consideration of the above, subject to the respective terms, conditions and obligations set forth herein, Worldwide Lint and ABCXYZ agree as follows:¶

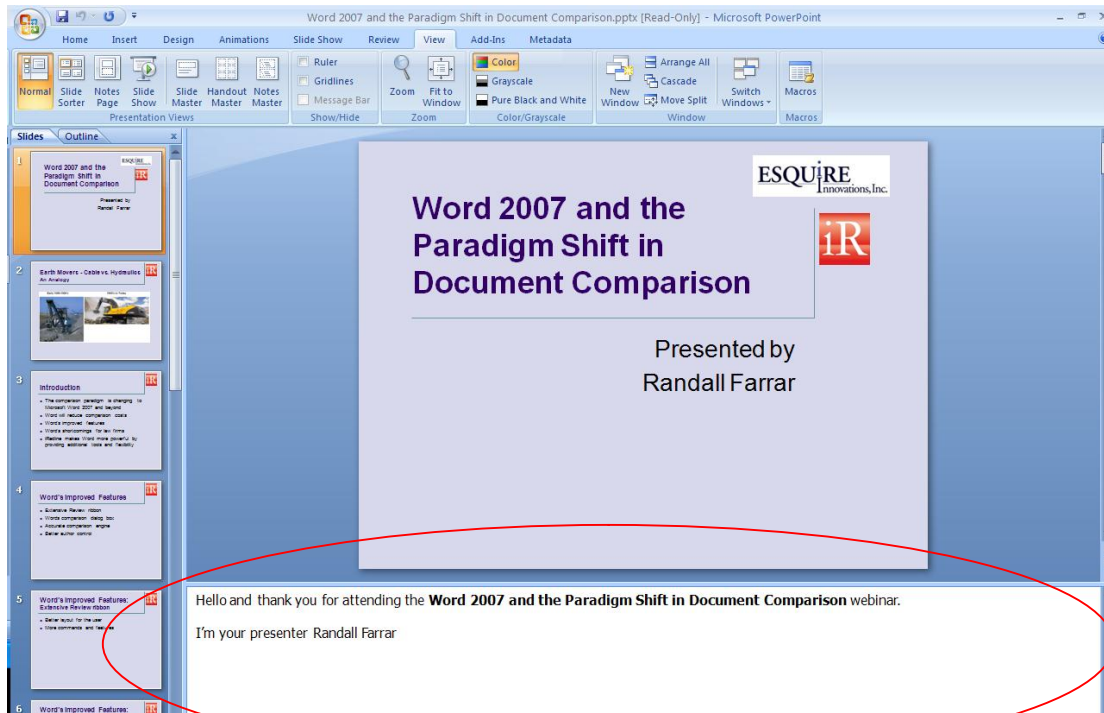
1. → Design of the Flash Calciner.¶

1.1. → Worldwide Lint's Design Costs. During the Term (as defined in Section 6 below) of this Agreement, Worldwide Lint will incur up to fifty thousand dollars (\$50,000) in expenses ("Worldwide Lint's Design Costs") paid to Coen Company, Inc., an international

Tracked Changes displayed in a Word Document

In Excel, metadata can also be very useful and can include formulas in a spreadsheet (a powerful feature for editing and calculating figures), hidden columns, author names and creation dates of documents.

In PowerPoint, metadata also includes author information and presentation creation dates as well as speaker notes and links to graphs or other statistics from outside documents.



PowerPoint showing Speaker notes

3. METADATA CAN BE HARMFUL

Metadata can be harmful when users unknowingly send out documents that contain confidential or potentially embarrassing information. There have been many well publicized cases where tracked changes or hidden comments have been left in a document sent via e-mail or shared on the internet:

Two examples of high profile metadata blunders can be found at the links below:

Hidden Text shows SCO Prepped a Lawsuit against BofA - http://news.cnet.com/2100-7344_3-5170073.html

A Microsoft Word document of SCO's suit against DaimlerChrysler originally identified Bank of America as the defendant instead of the automaker.

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.

http://www.timesonline.co.uk/tol/news/world/middle_east/article581486.ece

In Excel formulas can be useful, but if spreadsheets are linked to other workbooks, links should be disabled and formulas changes to static numbers before sending out, to prevent revealing information not intended for general distribution.

In PowerPoint Speaker notes are invaluable during a presentation, but something you will not want to share when sending a final copy of the presentation to attendees

4. METADATA CAN BE TRACKED CHANGES LEFT IN A DOCUMENT

The obvious and most publicized cases of document metadata involve tracked changes being left in a document sent out via e-mail or published on the internet.

This mistake can easily be made. 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 associate he is satisfied with the changes. He decides to send it to the client for review and clicks on the “Review” ribbon in Office 2007 and changes the document to “Final” in the “Tracking” section (see figure below)



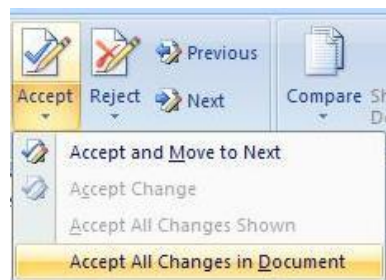
Review Ribbon in Word 2007

The tracked changes disappear from the document. He assumes they are no longer there, clicks on send via e-mail and sends the document to his client. The client opens the document to see all of the tracked changes displayed. This happened 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 his “Display for Review” settings were set by default to “Final Showing Markup”, thus revealing all of the changes in the document.

Track Changes Safe Guard Rules

If you want to make sure there are no tracked changes in your document, accept all changes just in case. This will give you peace of mind.

From the Review tab, you can click on the down arrow found on the Accept button and click Accept All Changes in Document. Whether you see the tracked changes or not, this will accept any stragglers.



Review Tab - Accept All Changes in Document

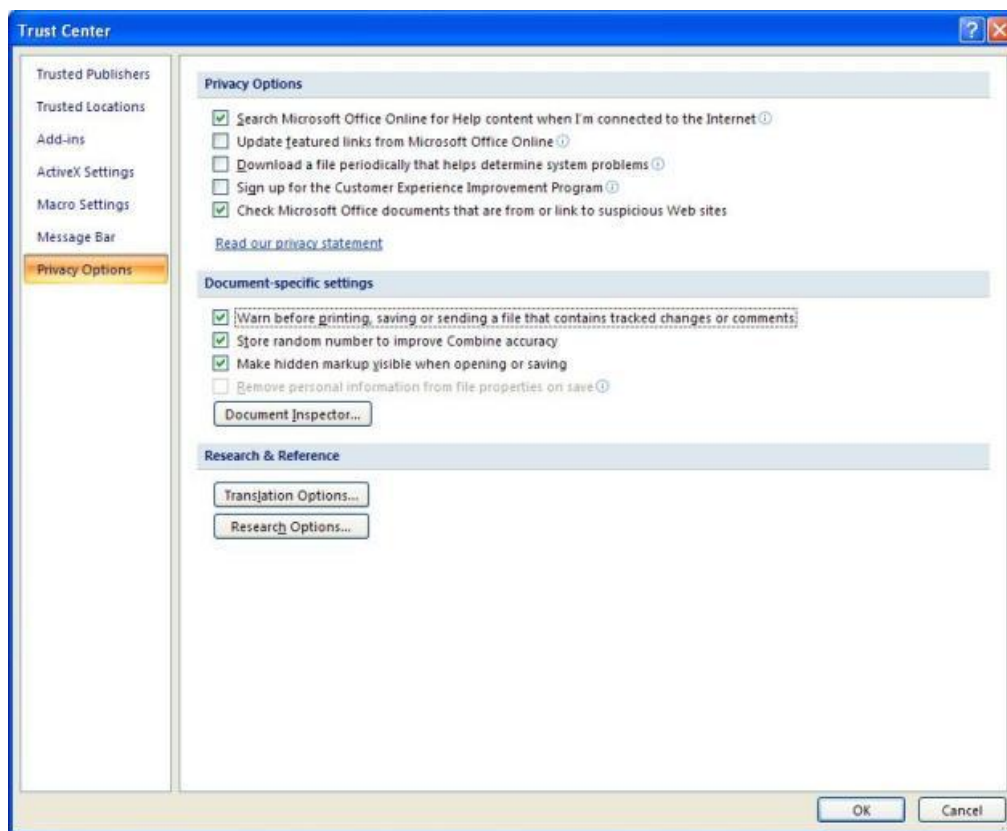
Turn on Word's Track Changes Warning System

There is an extra security feature in Word that warns you when you have tracked changes in your document when you print, save or email a document .

To turn Word's Track Changes warning system on, follow these steps:

- On the Office button select Word Options.
- In the Word Options dialog box select Trust Center.
- In the Options dialog box, select Security from the list on the left hand side, then click on the Trust Center Settings button.
- Under Document Specific Settings - Click the check box Warn before printing, saving or sending a file that contains tracked changes or comments.

(Note: if you firm uses a Metadata Software this feature should probably be turned off since it may interfere with the software, check with your vendor first).



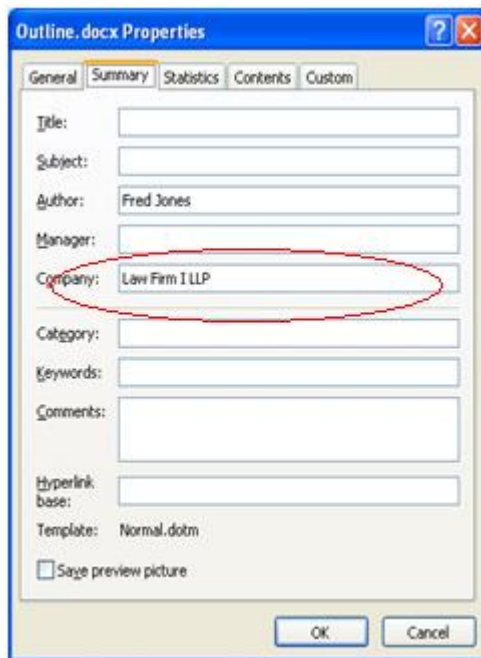
Privacy Options dialog

For added security, click the check box Make hidden markup visible when opening or saving. This will ensure that tracked changes are always visible.

5. METADATA CAN BE AUTHOR AND COMPANY INFORMATION

Did you know that 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 (found by accessing the Office Button then Word Options|Popular) for the last Saved by author, and will save the names if there have been multiple editors of a document (we tested adding up to 5,000 authors. Read about our test on www.randalfarrar.com). This is an example of application metadata.

If a document has been created from an earlier document using Save As, the author name of 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 it will stay with the document (see illustration below).



If a client billed for a document created by their highly paid attorney from ABC firm, sees a different author name and XYZ company listed in the properties it could lead to questions you don't want to answer.

Control Author Information

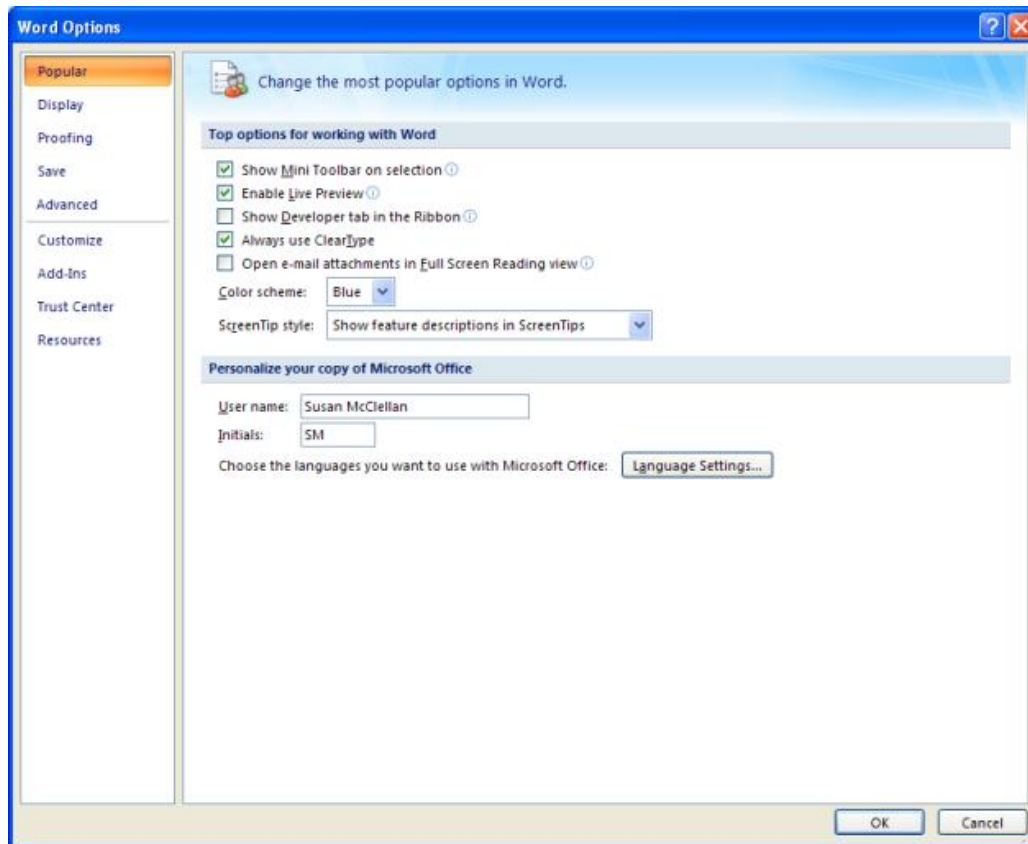
Microsoft Word has four areas that collect Author information:

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

User Name and User Initials Control it All

The User Name and User Initials control what ends up in the author properties of a Microsoft Word document.

User Name and User Initials are found in Word Option|Popular|Personalize your copy of Microsoft Office



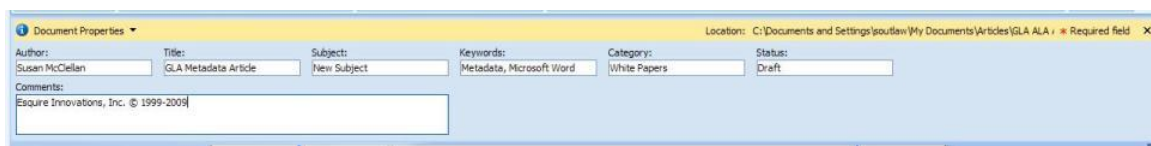
User Information settings

Unless you need the actual author name in this property for collaborative purposes, you may want to consider changing user information to something more generic .

Other Document Properties

Microsoft Word documents also contain other properties 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 your document.



The document author is pulled from the Word Options settings above, and is 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 you are using an automated be aware of these properties since they will remain with the document until they are changed or deleted.

Note: Dupe and revise (File...Save As) does change these properties.

In the second part of this series we will cover and other 5 Facts about Metadata. The more you know about potential security issues the easier it is to create and enforce a policy to manage these elements within your firm

Randall Farrar is the president and co-founder of Esquire Innovations, Inc. (www.esqinc.com), a software company that develops Microsoft Office integrated applications for the legal market, located in Temecula, CA. iScrub is their metadata management software used by over 450 law firms and corporations. He can be reached at randall.farrar@esqinc.com