
METADATA

UNCOVERING THE HIDDEN WORLD

Ron Wantland
April 17, 2007



METADATA

- **What** is it ?
- **Why** do we care ?
- **How** can we avoid problems ?

What is Metadata ?

- Data; more specifically data about data or data within data
- Hidden in Documents
- Generated during creation and editing
 - e.g., fragments of data from files that were previously deleted, overwritten, or worked simultaneously

Why Do We Care ?

- NY Bar Ethics Opinion, Dec. 8, 2004
 - Reasonable care to prevent disclosure of client confidences (<http://www.hricik.com/eethics/3.2.html>)
- Disclosure in Word document during run-up to Iraq war harmed reputation of Blair government
- Merck Pharmaceuticals inadvertently disclosed knowledge of Vioxx impact on heart attacks
- “Juicy” details of SCO draft complaint against Daimler-Chrysler released to public in Word
(http://news.com.com/2100-7344_3-5170073.html)

3 Types of Metadata in *Your* Files

- **Embedded Identifiers**
- **Hidden Text**
- **Track Changes**

Metadata and Common Applications

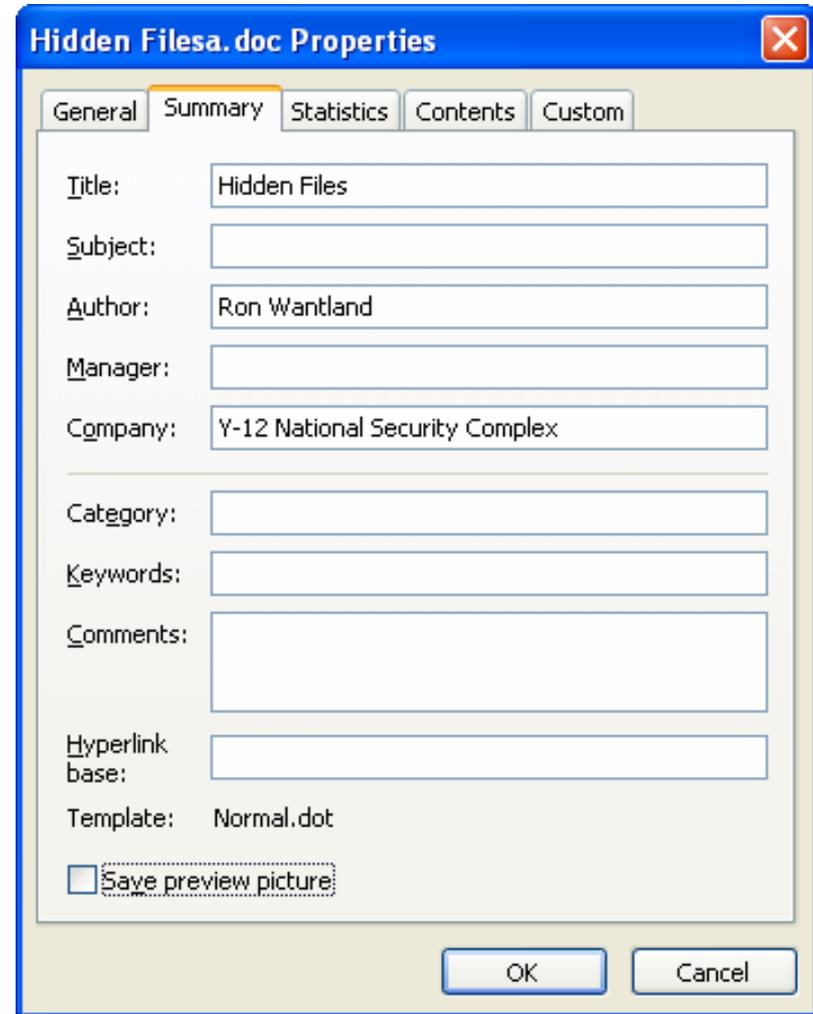
Metadata Type	Microsoft Word	Microsoft Excel	Microsoft PowerPoint
Embedded Identifiers	✓	✓	✓
Hidden Text	✓		
Track Changes	✓	✓	

Examples of Applications Yielding Metadata

- **Microsoft Office Products**
 - Excel, Word, Powerpoint, Vision, Access, Outlook, Publisher
- **RTF files**
- **HTML**
- **XML**
- **PDF (save as)**
- **Most software with UNDO function**
- **Office clones**
- **Wordperfect**
- **....., and more**

Embedded Identifiers in Microsoft Word

- **How do you know ...**
 - If there are embedded identifiers in your Microsoft Word document?
 - Open document
 - Click on File, Properties, Summary.
 - See if Author and Company names contain information.
- **How do you remove ...**
 - The identifiers?
 - Open document
 - Click on File, Properties, Summary.
 - Delete information in Author and Company fields and replace with spaces. Note that if you only delete the fields and do not replace them with spaces, it will pick up the default information again.



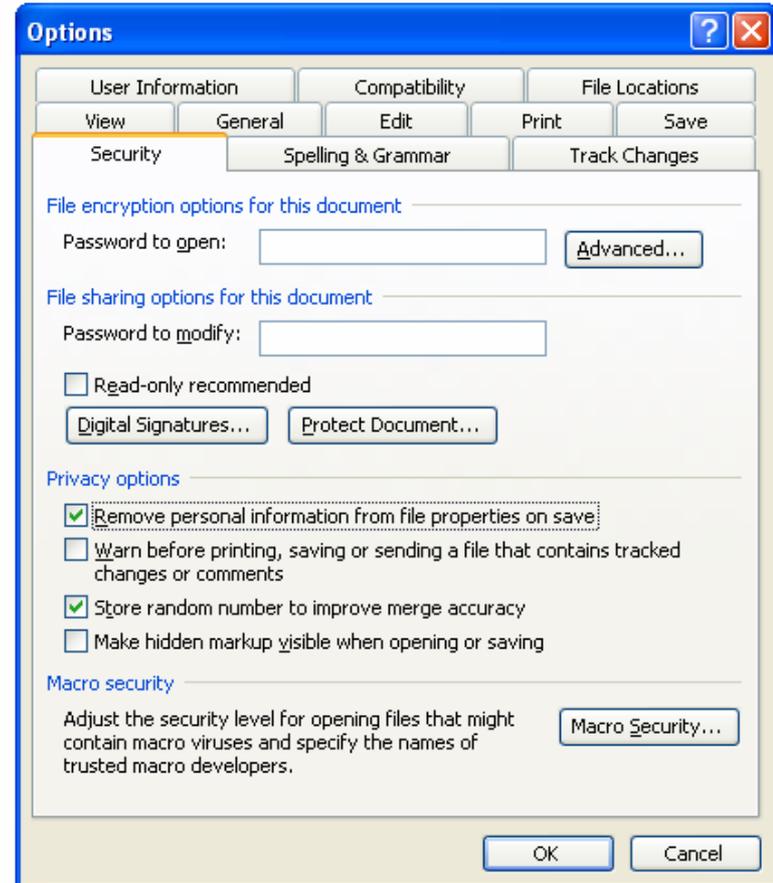
The screenshot shows the 'Hidden Files.doc Properties' dialog box with the 'Summary' tab selected. The fields are as follows:

Field	Value
Title:	Hidden Files
Subject:	
Author:	Ron Wantland
Manager:	
Company:	Y-12 National Security Complex
Category:	
Keywords:	
Comments:	
Hyperlink base:	
Template:	Normal.dot

At the bottom, there is a checkbox labeled 'Save preview picture:' which is currently unchecked. The 'OK' and 'Cancel' buttons are also visible.

Embedded Identifiers in Microsoft Word

- **How do you remove ...**
 - **Personal information from Microsoft Word documents?**
 - Open document.
 - Click Tools, Options, Security.
 - Look underneath Privacy Options section and check the first box that states “remove personal information from this file on save.”
 - Note that you have to do this for EVERY document that you open, this is not a global fix for every document.



Hidden Text in Microsoft Word

- **How do you know ...**
 - **If there is hidden text in your Microsoft Word doc?**
 - **Must show the Show/Hide (¶) command.**
 - **To enable this Show/Hide feature, simply click on the ¶ symbol on the Standard toolbar. This will show you all text (including headers/footers) and/or paragraph markings in the document.**

This is a test to show hidden data. It looks as though there is nothing else on this page but the text you see now.

|

¶
¶

This is a test to show hidden data. It looks as though there is nothing else on this page but the text you see now. ¶

THIS IS THE HIDDEN TEXT WITHIN THE FOOTER ¶

Track Changes in Microsoft Word

- **How do you know ...**
 - If there are tracked changes in a doc someone sends to you, or you are about to send out in a Microsoft Word doc?
- Viewed in Final Showing Markup
 - You can see tracked changes.
- Viewed in Final
 - You can't see tracked changes



This is a test for tracked changes.

This is a second test for tracked changes.

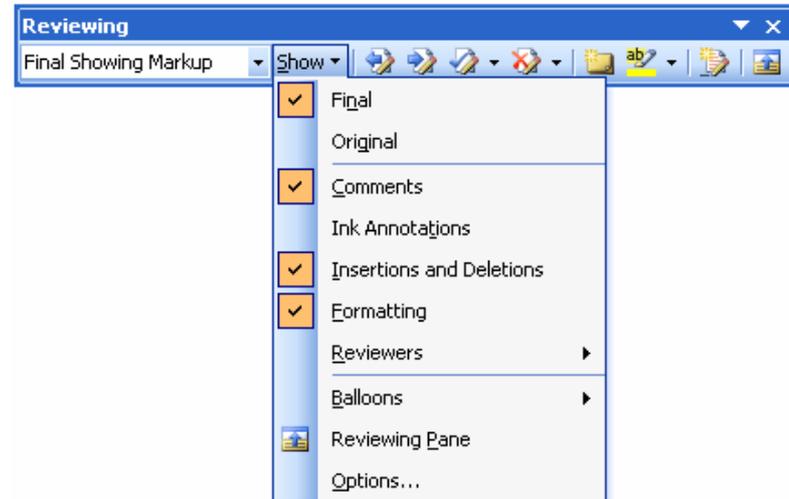


This is a test for tracked changes.

Here is another test for tracked changes.

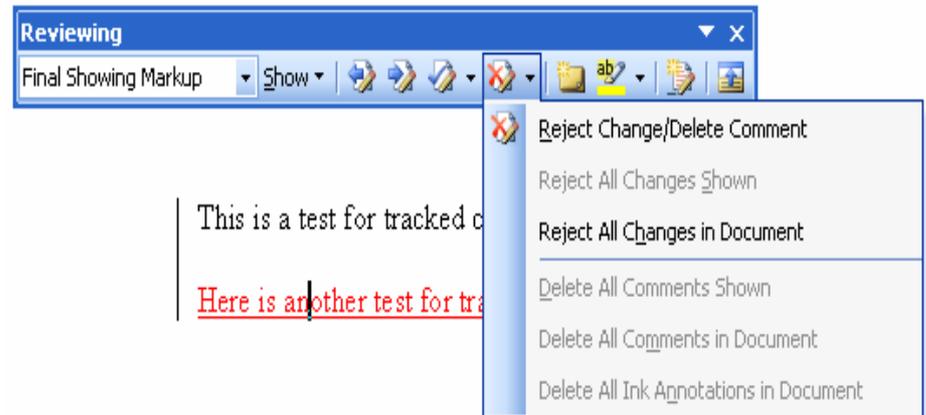
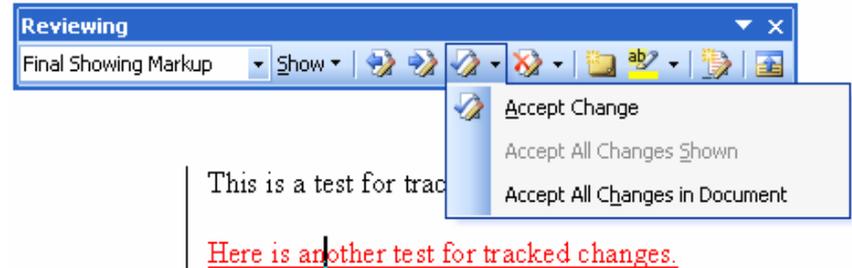
Track Changes in Microsoft Word (cont'd)

- **Are you sure you're seeing everything?**
 - Simply turning off the track changes button does **not** disable the feature.
 - Check all 3 blocks ...
 - the **Comments, Insertions and Deletions, and the Formats** to ensure you are seeing everything that has been done to the doc.



Track Changes in Microsoft Word (cont'd)

- **How do you remove ...**
 - All tracked changes before electronically sending the document out?
- **Must either accept or reject all changes in the finalized document**
 - One at a time, or
 - Globally.



Other Preventive Measures

- **Microsoft provides a free tool that purportedly removes metadata**
- **Commercial tools**
 - e.g., \$79 per workstation from Payne Consulting
(<http://www.payneconsulting.com/public/products/ProductDetail.asp?nProductID>)
- **Print, then fax, or scan as pdf file.**
 - **Caution: later versions of Adobe Acrobat are able to include comments, etc. from “save as” or “print as” options.**
- **Know your application software**

Y-12 OPSEC Initiatives

- **Conducted study on Personal Cell Phone Use**
- **Addressed Wireless Technology Issues**
- **Conducted Identity Theft Presentations**
- **Established “*Release of Public Information*” Review Process**
- **Conducted Internet Threat Vulnerabilities as local University**
- **Government Laptop Pool for Foreign Travel (ongoing)**
- **Gifts Received from Foreign Contacts Protocol**
- **“*The Insider Threat, Could It Be You?*”**