

# Guidelines for Editing Metadata

The Middle District of Pennsylvania is providing this guideline as a convenience to our efilers, and is unable to provide help-desk support for editing metadata. Efilers are responsible for ensuring that their documents are properly publicized by following the manuals and guides that pertain to the software that they are using.

# Table of Contents

Overview .....	2
What is metadata? .....	2
What are the issues associated with metadata? .....	2
What metadata is associated with a file? .....	3
Viewing Metadata .....	3
File Description Metadata .....	3
WordPerfect .....	3
Microsoft Word .....	3
Adobe Acrobat .....	5
Revision Metadata .....	5
WordPerfect .....	5
Microsoft Word .....	6
Adobe Acrobat .....	8
Removing Metadata .....	9
WordPerfect .....	10
Microsoft Word .....	11
Adobe Acrobat .....	13
Converting to PDF .....	16

## Overview

### What is metadata?

Metadata is data about data; hidden file information. Metadata becomes a problem when it is accidentally released. This document provides guidelines for removing metadata, or at least minimizing its exposure.

There are two types of metadata that are pertinent to us: *revision metadata* and *file description metadata*. *Revision metadata* helps with document editing. The user sees this data when reversing an edit or making other changes by using the **Undo** command. *Revision metadata* remains available until the document is saved using the **Save As**, **Print**, or **Convert (to PDF)** commands, but it is not necessarily removed when using the **Save** command. Documents are most susceptible to inadvertent metadata exposure during this phase.

*File description metadata* helps manage the file and includes document summary, headers, footers, hyperlinks, OLE object information, and routing slip information. Note: Both types of metadata vary among software applications.

### What are the issues associated with metadata?

Metadata can be accidentally released when the proper steps are not taken. This problem is predominantly associated with documents in WordPerfect or Microsoft Word that have simply been saved without taking any additional precautions such as removing hidden data, or disabling Track Changes in Microsoft Word and Save Undo/Redo Items in WordPerfect. Most word processing software stores up to the last 10 edits of a document as long as you continue to use **Save** instead of **Save As**.

Most word processing and publishing software provides tools to minimize the risk of inadvertently disclosing *revision metadata*. Alternatively, users should consider sharing documents in PDF format, as this greatly reduces the risk of exposure. Printing documents in PDF format removes *revision metadata*, but not all *file description metadata*. PDF files retain some basic *file description metadata*, such as author, file name, and date, which can be minimized if the proper conversion settings are used (see page 16).

Court users have expressed concerns regarding authorship metadata. When authors create a document using previous versions created by other authors, the original author's name is inherited, leading readers to think it was written by someone else. To avoid this issue, the user should verify and edit the author's name if incorrect through the Properties option (usually found under the File menu of word processing software, see *File Description Metadata* section on page 3). Otherwise, it is suggested to start with a fresh document, and copy and paste text from the existing documents.

There is also concern that some CM/ECF files may contain hidden metadata. The files in CM/ECF are in PDF format, so the risk of *revision metadata* being disclosed is minimal, and *file description metadata* can be minimized, as stated earlier. These final form documents are not where metadata or versioning data has been a problem.

## What metadata is associated with a file?

The type of metadata associated with a file depends on the word processing or publishing software. In the Administrative Office of the U.S. Courts and federal courts, Microsoft Word, WordPerfect, and Adobe Acrobat are most commonly used.

- WordPerfect metadata: comments, hidden text, annotations, undo/redo history, document summary data, headers, footers, hyperlinks, OLE object information, and routing slip information.
- Microsoft Word metadata: comments, revision marks from tracked changes, document version information, ink annotations, document properties, (including information from the Summary, Statistics, and custom tabs of the document Properties dialog box), email headers, routing slips, send-for-review information, document server properties, document management policy information, databinding link information for databound fields (last value will be converted to text), user name, template name, and text that is formatted as hidden (a font effect that is available in the font dialog box).
- PDF (Adobe Acrobat) metadata: embedded content; attached files; scripts; hidden layers; embedded search indexes; stored form data; review and comment data; hidden data from previous document saves; obscured text with images; comments hidden from the body of the PDF; unreferenced data, links, actions and JavaScript; and overlapping objects.

## Viewing Metadata

The steps for viewing metadata depend on the word processing or publishing software, and may even vary between software versions.

### File Description Metadata

Generally, *file description metadata* can be viewed under the Properties option of most software.

#### WordPerfect

Regardless of the version of WordPerfect, *file description metadata* can be viewed on the Summary tab of the Properties dialog box.

##### Viewing *file description metadata* in WordPerfect

1. Click **File** on the main menu bar.
2. Select **Properties**.
3. Select the **Summary tab**.
4. Click **OK**.

#### Microsoft Word

In Microsoft Word, this data can be viewed on several tabs of the Properties dialog box. The steps to access Properties are different for Word 2003, 2007, and 2010.

### Viewing file description metadata in Microsoft Word 2003

1. Click **File** on the main menu bar.
2. Select **Properties**.

### Viewing file description metadata in Microsoft Word 2007

1. Click the **Office** button on the main menu bar.
2. Select **Prepare**.
3. Click **Properties**. This will display the Document Information panel across the top of the document. To view the Document Properties dialog box, follow steps 4 and 5.
4. Click **Document Properties** in the Document Information panel.
5. Select **Advanced Properties** (Figure 1).

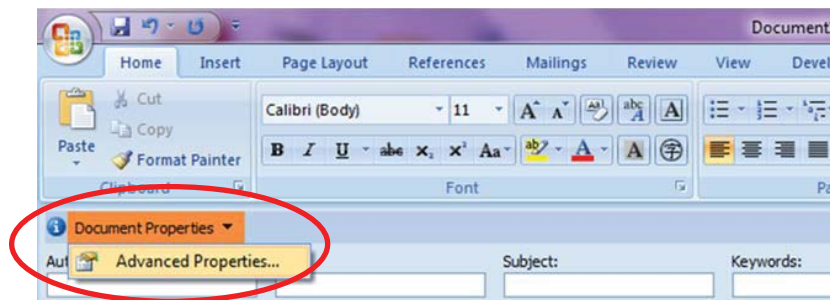


Figure 1: Viewing Document Properties dialog box (Word 2007)

### Viewing file description metadata in Microsoft Word 2010

1. Click the **File** button on the main menu bar.
2. Select **Info**. General document properties are displayed on the right side of the screen. To view the Document Properties dialog box, follow steps 3 and 4.
3. Click **Properties**.
4. Select **Advanced Properties** (Figure 2).



Figure 2: Viewing Document Properties dialog box (Word 2010)

## Adobe Acrobat

Regardless of the version of Adobe Acrobat, *file description metadata* can be viewed on the Description tab of the Properties dialog box.

### Viewing *file description metadata* in Adobe Acrobat

1. Click **File** on the main menu bar.
2. Select **Properties**.
3. Select the **Description** tab (default). To view additional *file description metadata*, click the **Additional Metadata** button.
4. Click **OK** on Additional Metadata dialog box, if opened.
5. Click **OK** on Properties dialog box.

## Revision Metadata

The methods of viewing of *revision metadata* vary between word processing and publishing software, as well as between versions.

### WordPerfect

The file format of WordPerfect makes it easy to view *revision metadata* by using the reveal codes feature (Figure 3).

### Viewing *revision metadata* in WordPerfect

1. Click **View** on the main menu bar.
  2. Select **Reveal Codes**.
- OR
- Type **Alt + F3**.

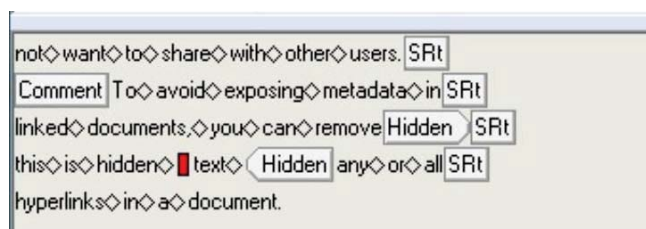


Figure 3: WordPerfect's Reveal Codes showing document metadata (comments and hidden text)

## Microsoft Word

In Microsoft Word, most *revision metadata* can be accessed by having the proper security settings. The steps needed to set proper security settings vary for Word 2003, 2007, and 2010.

### Viewing *revision metadata* in Microsoft Word 2003

1. Click **Tools** on the main menu bar.
2. Select **Options**.
3. Select the **Security** tab in the Options dialog box.
4. Check **Warn before printing, saving or sending a file that contains tracked changes or comments** and **Make hidden markup visible when opening or saving**, if they are unchecked (Figure 4).
5. Click **OK**.

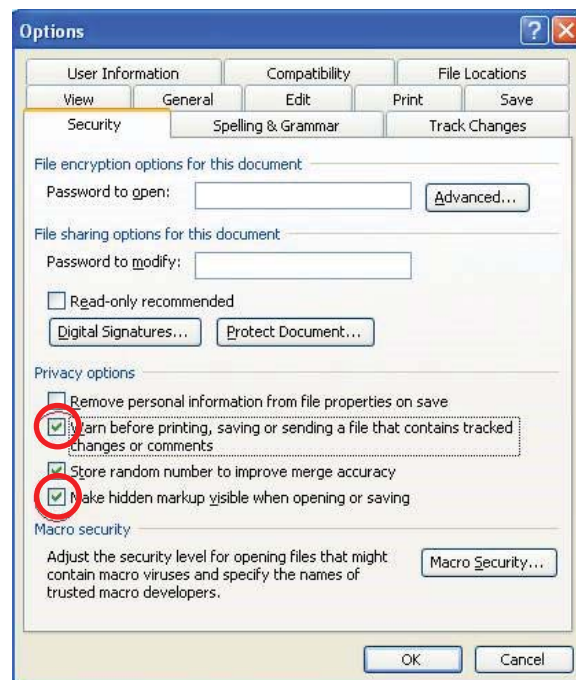


Figure 4: Security Settings for displaying revision metadata (Word 2003)

### Viewing *revision metadata* in Microsoft Word 2007

1. Click the **Office** button on the main menu bar.
2. Click **Word Options** at the bottom right of the menu; this will display the Word Options dialog box.
3. Select the **Trust Center** option on the left pane.
4. Click the **Trust Center Settings** button; this will open the Trust Center dialog box.
5. Select **Privacy Options** in the left pane.
6. Check **Warn before printing, saving or sending a file that contains tracked changes or comments** and **Make hidden markup visible when opening or saving**, if they are unchecked (Figure 5).
7. Click **OK**.

### Viewing *revision metadata* in Microsoft Word 2010

1. Click the **File** button on the main menu bar.
2. Select **Options**; this will display the Word Options dialog box.
3. Select the **Trust Center** option on the left pane.
4. Click the **Trust Center Settings** button; this will open the Trust Center dialog box.
5. Select **Privacy Options** in the left pane.
6. Check **Warn before printing, saving or sending a file that contains tracked changes or comments** and **Make hidden markup visible when opening or saving**, if they are unchecked (Figure 5).
7. Click **OK**.

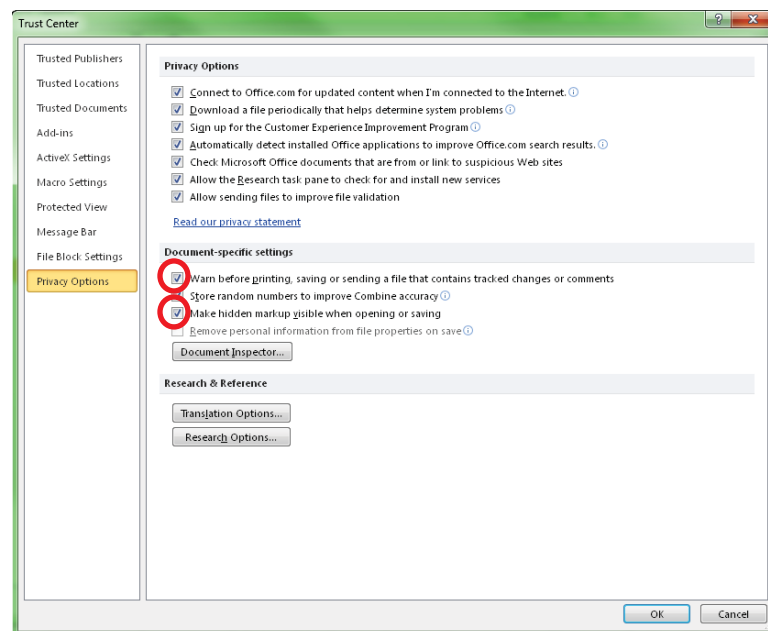


Figure 5: Privacy Options for displaying revision metadata (Word 2010)



## Adobe Acrobat

The Examine Document tool in Adobe Acrobat 8 and 9 scans the PDF and identifies potentially hidden information: metadata, comments, bookmarks, file attachments, and even “hidden text” (text hidden by another object or white text on white background). In Adobe Acrobat X and Adobe Acrobat X Pro, the tool is now called Remove Hidden Information.



Please note that “hidden text” is not only an issue of inadvertently releasing metadata but also of improper redaction practices.<sup>1</sup>

### Viewing *revision metadata* in Adobe Acrobat 9

1. Click **Document** on the main menu bar.
2. Click **Examine Document**; this will display results on the Examine Document pane on the left side of the screen.
3. Click the **expand** icon on the Examine Document pane so that all information is displayed on the results list. To view metadata, click the **Show preview** option (Figure 6).
4. Close the Examine Document pane.

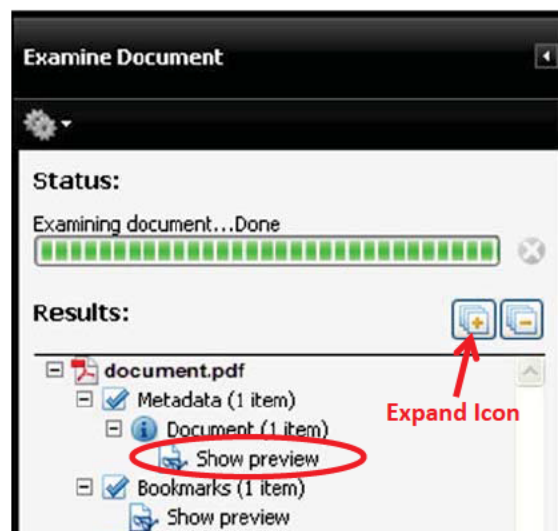


Figure 6: Examine Document (Acrobat 9)

<sup>1</sup> Additional information regarding this issue as it relates to the courts can be found on the Utah District Court blog entry “Redaction Warnings!” published online at <http://utd-cmecf.blogspot.com/search?q=metadata>.

### Viewing *revision metadata* in Adobe Acrobat X and Adobe Acrobat X Pro

1. Click **Tools** on the top right of the screen.
2. Select **Protection**.
3. Select **Remove Hidden Information**; this will open the Remove Hidden Information pane on the left side of the screen.
4. Click the **expand** icon on the Remove Hidden Information pane so that all information is displayed on the results list. To view metadata, click the **Show preview** option (Figure 7).
5. Close the Remove Hidden Information pane.

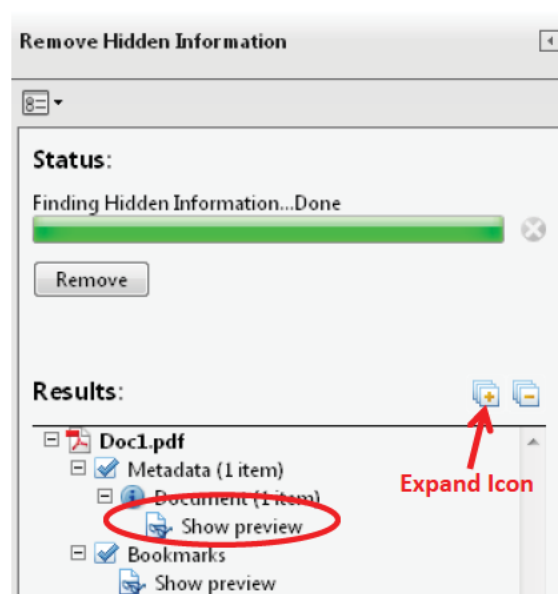


Figure 7: Remove Hidden Information Pane (Acrobat X Pro)

## Removing Metadata

Depending on the software, there are several methods to remove metadata. In most word processing applications, **Save As** creates a clean file, removing *revision metadata* (results may vary depending on version). You can also use the PDF printer (Distiller) to create a clean file. In addition, there are metadata removal tools available through word processing and publishing software or third parties.

## WordPerfect

Steps and tools for removing metadata in WordPerfect depend on the software version.

### Removing metadata in WordPerfect 12 or earlier

You will need to manually remove the *revision* and *file description* metadata. The link below provides instructions on how to manually clean up metadata.

[Minimizing metadata in WordPerfect 12 document](#)<sup>2</sup>

### Removing metadata in WordPerfect X3 or newer

1. Click **File** on the main menu bar.
2. Select **Save Without Metadata**.
3. Check the metadata that you wish to remove.
4. Click **Save**.

The link below provides additional information on using the Save Without Metadata option.

[Saving documents without metadata](#)<sup>3</sup>



**Note**, if you are using WordPerfect X3 or newer and do not see the Save Without Metadata option, you can add it to your menu bar.

### Adding the Save Without Metadata option to WordPerfect X5

1. Click **Tools** on the main menu bar OR right-click on the main menu bar and select Settings. (If you do this, skip to number 4.)
2. Select **Settings** from the dropdown list.
3. Click **Customize** on the Settings dialog box.
4. Select the **Menus** tab on the Customize Settings dialog box.
5. Click the **Edit** button.
6. Under Features, select **Save Without Metadata** and drag it to **File** on the main menu bar (Figure 8). Or click **Add Menu Item**, which adds the option directly to the main menu bar.
7. Click **OK** in the Menu Editor dialog box.
8. Click **Close** in the Customize Settings dialog box.
9. Click **Close** in the Settings dialog box, if it is open.

<sup>2</sup> [http://www.corel.com/content/pdf/wpo12/Minimizing\\_Metadata\\_In\\_WordPerfect12.pdf](http://www.corel.com/content/pdf/wpo12/Minimizing_Metadata_In_WordPerfect12.pdf)

<sup>3</sup> <http://www.corel.com/corel/pages/index.jsp?pgid=800411&item=resource&listid=1600090>

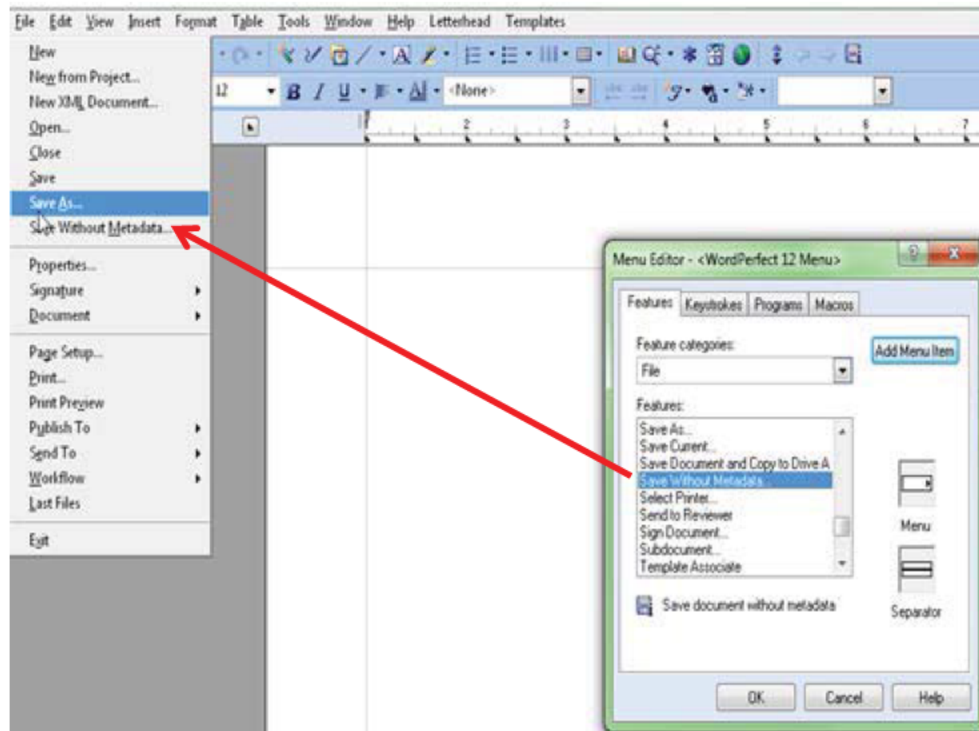


Figure 8: Adding Save Without Metadata to File menu

## Microsoft Word

In Word 2003, you must remove metadata manually. However, in Word 2007 or 2010, the Document Inspector feature (Figure 9) allows you to select the metadata you want to remove.

### Removing metadata in Word 2003

The link below provides instructions on how to manually clean up *revision* and *file description* metadata.

[How to minimize metadata in Word 2003](http://support.microsoft.com/kb/825576)<sup>4</sup>

<sup>4</sup> <http://support.microsoft.com/kb/825576>

### Removing metadata in Microsoft Word 2007

1. Click the **Office** button on the main menu bar.
2. Select **Prepare**.
3. Click **Inspect Document**; this will display Document Inspector dialog box (Figure 9).
4. Check the content you want to inspect for.
5. Click **Inspect**.
6. Click the **Remove All** button next to the content you would like to remove.
7. Click **Close**. If you would like to reinspect the document to verify content was removed, click **Reinspect**; this will take you back to step 4.

Additional information may be found using the following link:

[Remove hidden data and personal information from Office documents](http://office.microsoft.com/en-us/excel-help/remove-hidden-data-and-personal-information-from-office-documents-HA010037593.aspx)<sup>5</sup>

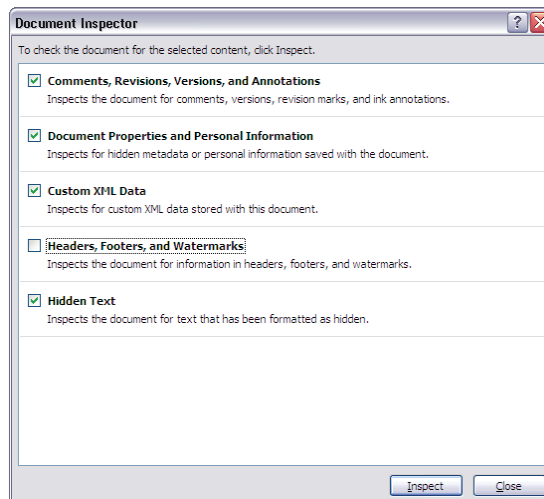


Figure 9: Document Inspector dialog box for reviewing metadata (Word 2007)

<sup>5</sup> <http://office.microsoft.com/en-us/excel-help/remove-hidden-data-and-personal-information-from-office-documents-HA010037593.aspx>

### Removing metadata in Microsoft Word 2010

1. Click the **File** button on the main menu bar.
2. Select **Info**.
3. Click **Check for Issues**.
4. Select **Inspect Document**; this will display Document Inspector dialog box (Figure 10).
5. Check the content you want to inspect for.
6. Click **Inspect**.
7. Click the **Remove All button** next to the content you would like to remove.
8. Click **Close**. If you would like to reinspect the document to verify that content was removed, you can do so by clicking **Reinspect**; this will take you back to step 5.

Additional information may be found using the following link:

[Remove hidden data and personal information by inspecting documents Word 2010](http://office.microsoft.com/en-us/word-help/remove-hidden-data-and-personal-information-by-inspecting-documents-Word-2010)<sup>6</sup>

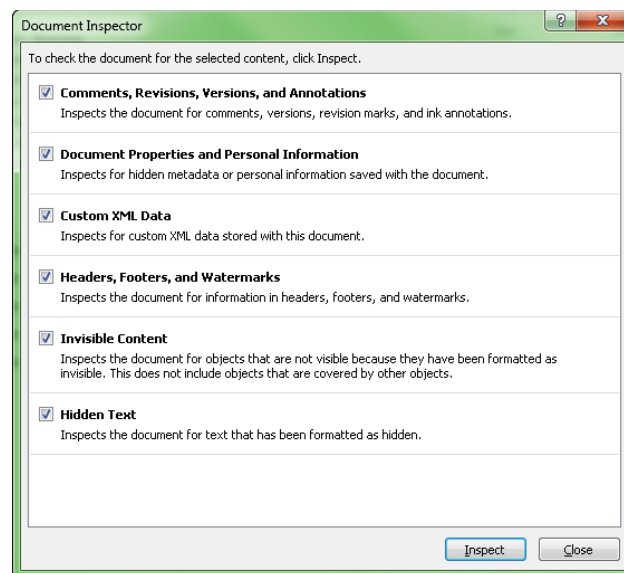


Figure 10: Document Inspector dialog box for reviewing metadata (Word 2010)

## Adobe Acrobat

As mentioned previously, Adobe Acrobat 8 and 9 provide the Examine Document function, which removes metadata. This feature scans the document and allows you to remove any hidden information with a single click. In Adobe Acrobat X and Adobe Acrobat X Pro, the tool is called Remove Hidden Information. Adobe Acrobat X Pro's Sanitize Document option strips the document by converting it into an image; however, this also removes the ability to perform text searches, so the Sanitize Document option is not recommended.

<sup>6</sup> <http://office.microsoft.com/en-us/word-help/remove-hidden-data-and-personal-information-by-inspecting-documents-HA010354329.aspx>

### Removing metadata in Adobe Acrobat 9

1. Click **Document** on the main menu bar.
2. Click **Examine Document**; this will display results on the left-side navigation pane.
3. Uncheck the content or metadata that you do not want to remove (Figure 11).
4. Click **Remove**.
5. Click **OK**.

To configure Adobe Acrobat to remind you to remove metadata when closing, follow the steps below:

1. Click **Edit** on the main menu bar.
2. Click on **Preferences**.
3. Select **Documents** from the Categories list on the left pane.
4. Check **Examine document when closing document**.
5. Click **OK**.

Additional information may be found using the following link:

[Examine a PDF for hidden content – Adobe Acrobat 9 Standard Help](http://help.adobe.com/archive/en_US/acrobat/9/standard/acrobat_standard_9.0_help.pdf#page=242)<sup>7</sup>

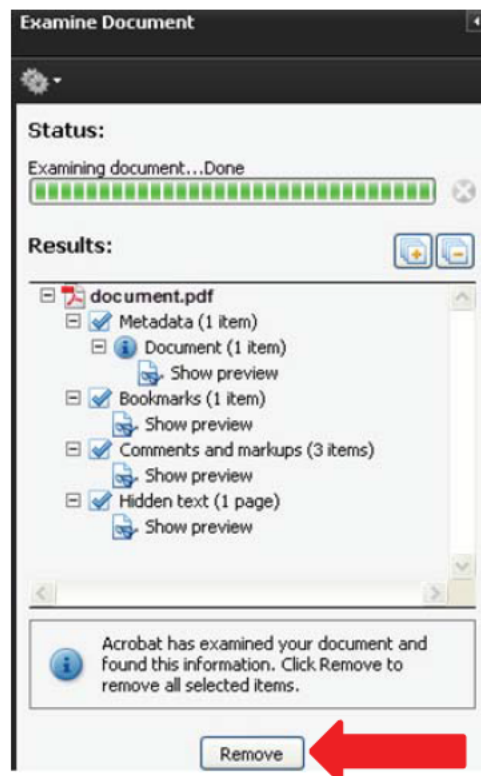


Figure 11: Examine Document (Acrobat 9)

<sup>7</sup> [http://help.adobe.com/archive/en\\_US/acrobat/9/standard/acrobat\\_standard\\_9.0\\_help.pdf#page=242](http://help.adobe.com/archive/en_US/acrobat/9/standard/acrobat_standard_9.0_help.pdf#page=242)

### Removing metadata in Adobe Acrobat X and Adobe Acrobat X Pro

1. Click **Tools** on the top-right of the screen.
2. Select **Protection**.
3. Select **Remove Hidden Information**; this will open the Remove Hidden Information pane on the left side of the screen.
4. Uncheck the content or metadata that you do not want to remove. This allows you to sanitize the document without removing the text search capability. In order to preserve text search capabilities, you should uncheck overlapping objects (Figure 12).
5. Click **Remove**.
6. Click **OK**.

To configure Adobe Acrobat to remind you to remove metadata when closing , follow the steps below:

1. Click **Edit** on the main menu bar.
2. Click on **Preferences**.
3. Select **Documents** from the Categories list on the left pane.
4. Check **Remove Hidden Information when closing document**.
5. Click **OK**.

Additional information may be found using the following link:

[Remove Sensitive Content \(Metadata and Redaction\) – Adobe Acrobat X Pro Help](http://help.adobe.com/en_US/acrobat/X/pro/using/WS4E397D8A-B438-4b93-BB5F-E3161811C9C0.w.html)<sup>8</sup>

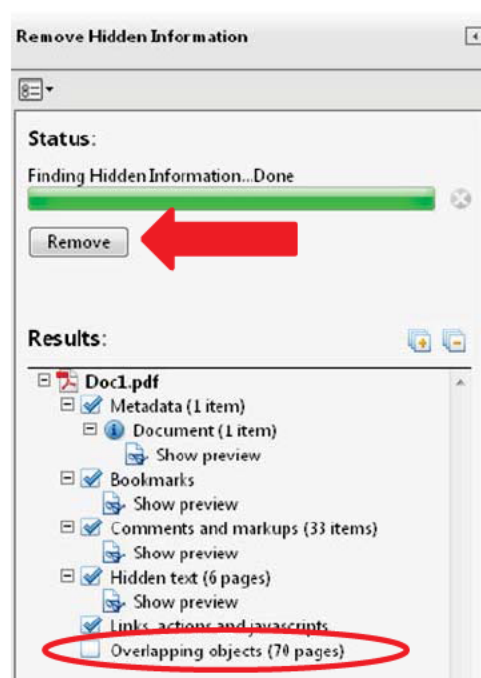


Figure 12: Remove Hidden Information Pane (Acrobat X Pro)

<sup>8</sup> [http://help.adobe.com/en\\_US/acrobat/X/pro/using/WS4E397D8A-B438-4b93-BB5F-E3161811C9C0.w.html](http://help.adobe.com/en_US/acrobat/X/pro/using/WS4E397D8A-B438-4b93-BB5F-E3161811C9C0.w.html)



Adobe Acrobat also offers ways to automate metadata scrubbing. We do not recommend any method in particular, as this depends on user need. Scrubbing can be a backend solution performed by the IT staff (see references listed below). Additionally, Adobe Acrobat X Pro provides a new feature (Action Wizard) that can be used to automate the metadata removal.<sup>9</sup>

Information on how to automate metadata scrubbing:

[Scrubbing Metadata – Practice and Policy](#)<sup>10</sup>

[Find and Remove PDF Metadata](#)<sup>11</sup>

## Converting to PDF

Converting documents from word processing to PDF format scrubs their *revision metadata*. Prior to converting or printing to PDF, the document should be properly scrubbed using tools provided by the authoring word processor or other tools available. In addition, you should verify that the proper conversion/print settings are configured. Some of the inherited *file description metadata* includes document author, document title, document keyword, etc.

### Setting Adobe PDF Conversion Settings (Print to PDF – Microsoft Word 2010/WordPerfect)

1. Click **File** on the main menu bar.
2. Select **Print**.
3. Select **Adobe PDF** as printer on the Print dialog box.
4. Click **Printer Properties** (Word) or **Properties** (WordPerfect); this will open the Adobe PDF Document Properties dialog box.
5. Uncheck the **Add document information** checkbox (Figure 13).
6. Click **OK**.

<sup>9</sup> The following link provides additional guidance for Action Wizard: <http://tv.adobe.com/watch/acrobat-x/how-to-create-actions/>.

<sup>10</sup> <http://blogs.adobe.com/acrobat/scrubbing-metadata-practice-and-policy/>

<sup>11</sup> <http://www.microsystems.com/resources/wordtips/wordtip003.php>

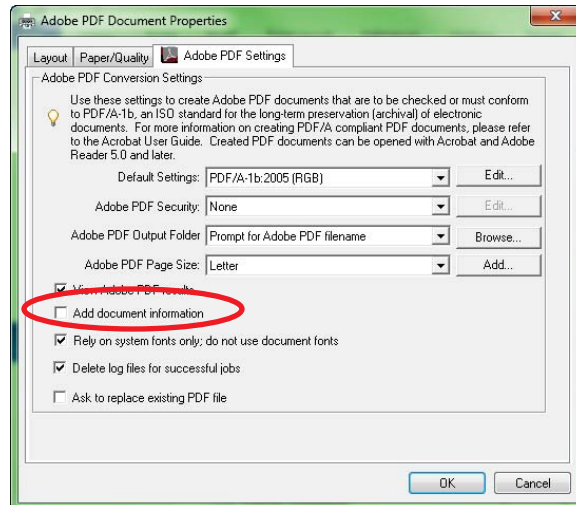


Figure 13: PDF Settings when printing to PDF file

### Setting Adobe PDF Conversion Settings (Save as Adobe PDF – Microsoft Word 2010)

1. Click **File** on the main menu bar.
2. Select **Save as Adobe PDF**; this will open the Save Adobe PDF File As dialog box.
3. Click the **Options** button; this will open the Acrobat PDFMaker dialog box.
4. Uncheck the **Convert document information** checkbox (Figure 14).
5. Click **OK**.
6. Click **Save**.

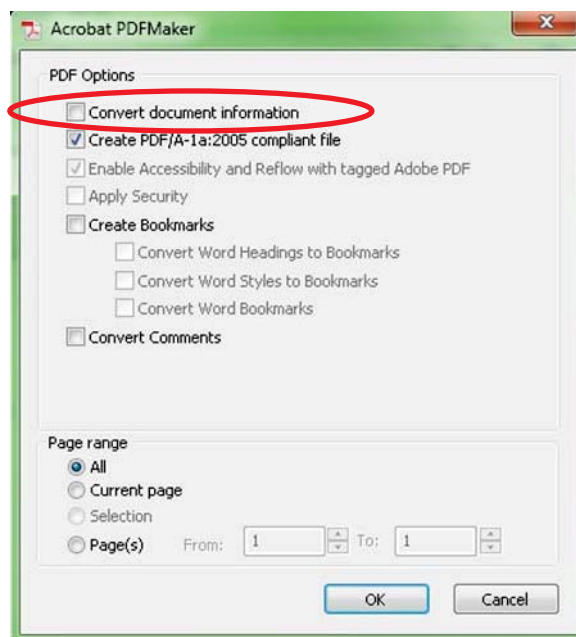


Figure 14: Adobe PDFMaker dialog box (Word 2010)