



Metadata: Check It, Scrub It, and Control What Your Files Reveal

The hidden data inside every document, and why it matters most in discovery

A lawyer produces discovery as a Word file. Opposing counsel opens it, makes four edits, and sends it back as if that were the operative version. Set the gamesmanship aside; the deeper problem is what the original file carried out the door. Tracked changes, comments, the names of everyone who touched it, when they touched it, and sometimes earlier drafts of the very language being produced. Metadata is the data inside your document about your document, and left unchecked it tells the other side things you never meant to say.

What metadata is

Every Office file stores hidden information alongside the visible content: author and company name, total editing time, the template it came from, tracked changes and comments (even when hidden from view), hidden text, earlier versions, and custom properties. Photos carry it too. None of it shows on the page; all of it ships with the file.

Check it (Word, Excel, PowerPoint)

File > Info. The Properties panel on the right shows author, dates, and more at a glance. For the full sweep, File > Info > Check for Issues > Inspect Document. The Document Inspector scans for comments and revisions, document properties and personal information, hidden text, headers and footers, and more, then reports what it finds.

Scrub it

In the Document Inspector results, click "Remove All" next to each category you want gone: comments and revisions, personal information, hidden text. Do this before the file leaves your office.

Two things the inspector will not decide for you, so handle them first:

- Accept or reject all tracked changes yourself (Review > Accept > Accept All Changes), then turn Track Changes off. You want the text settled deliberately, not silently flattened.
- Read and delete any comments you still need to act on before they disappear.

The stronger move for anything you produce: convert to PDF

For discovery and anything going to an adversary, produce a PDF, not a live Word file. A PDF is not editable in the casual way a .docx is, and it does not carry tracked changes or comments as working layers.

But a PDF is not automatically clean. "Save as PDF" from Word can carry document properties, and PDFs keep their own metadata (author, title, the software that produced them). So:

- Check a PDF's metadata in your PDF tool; in Acrobat, File > Properties.
- Scrub it: Acrobat's "Sanitize Document" or "Remove Hidden Information" strips metadata, hidden layers, and embedded data.
- If you live in a terminal, exiftool reads and strips metadata from PDFs and images on any platform:

```
exiftool document.pdf      # show all metadata
exiftool -all= document.pdf # strip all metadata (writes a cleaned copy)
```

A word on redaction

Drawing a black box over text, or highlighting it black, hides nothing; the text sits underneath and copies straight out. True redaction removes the underlying content. Use a real redaction tool, run the sanitize step above, then verify by trying to select the "redacted" text. If it selects, it is not redacted.

Add the metadata you want

Metadata is not only a liability; you can set it on purpose. File > Info > Properties lets you set Title, Tags, and Author. Set the author to your firm name instead of whichever staff member happened to create the file, add the tags your document system indexes on, and give it a clean title. Then the metadata that does ship says exactly what you want it to.

Why this matters for a legal practice

Documents you produce, file, or send to an adversary get inspected by people looking for advantage. Tracked changes reveal your negotiating history. Comments reveal your internal analysis. Author and timestamp data reveal who did what and when. Making metadata inspection a fixed step before any document leaves the office, and producing in sanitized PDF, is basic professional hygiene that has saved lawyers from real embarrassment and worse.

The checklist before you send

- Accept or reject tracked changes; turn Track Changes off.
- Delete comments.
- Run Document Inspector; Remove All on properties, personal information, hidden text.
- Convert to PDF for production.

- Sanitize the PDF; verify redactions by trying to select the text.
- Keep your own copy of exactly what you sent.

Using AI to help with this. Use it as a preflight reviewer: "List everything in a Word document that could leak metadata to opposing counsel in discovery, and give me the exact click path in current Word to find and remove each one before I convert to PDF." Then follow the steps yourself and verify.