NYPRR | June 2007

Metadata and Lawyers - an Issue That Will Not Die - Part I

BY LAZAR EMANUEL

[Author's note: I am indebted for much of the technical material in this article to my son David, IT consultant, who has guided me, as he always does, through the maze of technology surrounding computers.]

A March 14 Opinion of the Alabama State Bar Disciplinary Commission confirms that the issues created for lawyers by metadata remain essentially unresolved. The Alabama opinion cites and relies upon two Ethics Opinions of the New York State Bar and ignores an essentially contrary Opinion of the ABA. Recent opinions from Florida and Maryland have added to the confusion. [*See*, Florida, Opinion 06-02; Maryland, Opinion 2007-09.]

The New York view is two-pronged: (1) a lawyer has a duty to purge from an electronic document he intends to send to the opposing attorney all metadata which might disclose a client's secrets or confidences (NYSBA Opinion 782); and (2) a lawyer who receives an electronic document from his adversary may not make use of computer software applications to "mine" (i.e., search for metadata in) the document (NYSBA Opinion 749). The ABA, on the other hand, concludes as follows [Formal Opinion 06-442 (August 5, 2006)]:

> This opinion addresses whether the ABA Model Rules of Professional Conduct permit a lawyer to review and use embedded information contained in e-mail and other electronic documents, whether received from opposing counsel, an adverse party, or an agent of an adverse party. The Committee concludes that the Rules generally permit a lawyer to do so.

The U.S. Supreme Court has compounded the confusion by its amendments to Federal Civil Procedure Rule 26, which became effective on December 1, 2006. The amendments require the parties to litigation to discuss discovery of electronically stored information (ESI) during the discovery process at the beginning of the litigation. The Rule anticipates that the parties will make all relevant ESI available for discovery, except for ESI that a party identifies as "not reasonably accessible because of undue burden or cost." The import of this language is to make a party's relevant metadata potentially available for discovery.

Neither New York nor the ABA has addressed the issue of metadata and discovery. Apparently because it was written after the amendments to the Federal Rules, the Alabama opinion, *supra*, does:

> One possible exception to the prohibition against the mining of metadata involves electronic discovery. Recent court decisions indicate that parties may be sanctioned for failing to provide metadata along with electronic discovery submissions. In certain cases, metadata evidence may be relevant and material to the issues at hand... In Enron type litigation, the mining of metadata may be a valuable tool in tracking the history of accounting decisions and financial transactions.
>
> The production of metadata during discovery will ordinarily be a legal matter within the sole discretion

> of the courts. The Commission advises attorneys, however, to be cognizant of the issue of disclosing metadata during discovery. Both parties should seek direction from the court in determining whether a document's metadata is to be produced during discovery (Office of Alabama General Counsel, Ethics Opinion RO-2007-02 (3/14/07)).

At this juncture, therefore, we are left with the following issues:

1. Does a lawyer who sends a document electronically have an obligation to exercise reasonable care to ensure that the document does not inadvertently disclose a secret or confidence of her client through metadata? If so, what constitutes "reasonable care"?

2. Does a lawyer who receives an electronic document from opposing counsel have an obligation: (a) to refrain from researching, reviewing and using metadata embedded in the document; and/or (b) to notify the sender that he knows or suspects that the document contains metadata which she appears to have forwarded inadvertently?

3. What is the impact of amended Federal Rule 26 on these issues?

I will reserve these issues for discussion in Part II of this article. In this Part I, I will offer a definition of metadata and information about a few tools for identifying and controlling it.

Definition of Metadata.

Surprisingly, the term "metadata" does not appear in the Eleventh Edition (2003) of Merriam-Webster's Collegiate Dictionary. In ancient Greece, "meta" was the name for the turning point or winning post in a race-course, and it came to designate a goal or target. In modern usage, it is generally the first part of a word that gives the second part a sense of extension, inclusion or enlargement (e.g., meta-galaxy = the entire system of galaxies; meta-ethics = the study of the meaning of ethical terms and concepts).

In the word "metadata," the term "meta" connotes the data in an electronic file (the electronic record behind a finished document). It is generally referred to as "data about data". As ABA Formal Opinion 06-442 (August 5, 2006) advises: "E-mail and other electronic documents often contain 'embedded' information. Such embedded information is commonly referred to as metadata."

A more extended definition of the term is supplied by the ABA publication, *"What's the Meta with Metadata?"* by Peter H. Geraghty, Director, ABA Ethic Search (e-news for members, January 2006):

> Metadata may reveal who worked on a document, the name of the organization that created or worked on it, information about prior versions of the document, recent revisions, and comments inserted in the document during drafting or editing.... The hidden text may reflect editorial comments, strategy considerations, legal issues raised by the client or the lawyer, or legal advice provided by the lawyer... *ABA/BNA Lawyers' Manual on Professional Conduct* 21 Current Rep. 39 (2004).

An even more detailed discussion is contained in the article, *"Transmission and Receipt of Invisible Confidential Information,"* by David Hricik and Robert Jueneman, 15 No. 1 Prof. Law 18 (2004), which lists all the following as potential metadata: author's name or initials; the names of previous authors; firm

name; computer name; name of the network server or hard disk where the document is stored; document revisions (e.g., Microsoft Track Changes); previous versions; template information; hidden text; and comments. (My own experience has taught me that in Microsoft Word, Track Changes and comments are especially potent and potentially dangerous sources of metadata.)

Inherent Metadata Risks

The risk for a lawyer is that, however careful he may have been, a document transmitted by him electronically is apt to contain metadata which is inimical to his client's interest or, in the worst case, reveals the client's secrets or confidences. My limited research leads me to conclude that many solo practitioners and many lawyers in small or medium-sized firms don't know what metadata is or don't understand its potential for damage.

In the process of preparing a final document, a lawyer who uses all the word- or data-processing tools on his computer will have gone through several steps, all of them ostensibly buried in the document and invisible to anyone but the lawyer himself. The reality, however, is that the document's history is embedded in his files and is in fact available to anyone -including opposing counsel - who receives the document electronically.

Here's one example of metadata's potential impact. A partner in a law firm, too busy to write a memo of law requested by his client, assigns the memo to his associate. The associate writes the memo and the partner reviews it. With minor changes, the memo goes out to the client, together with the firm's bill, reciting the partner's hourly rate. Buried in the document, traceable by the client, is the name of the original author of the memo - the associate - as well as the time spent by each lawyer on the memo.

Here's another example from my own recent experience. I was retained by a young doctor educated abroad to review an employment contract tendered to him by a New York hospital. The contract was embodied in a standard form, modified in several respects specifically for my client. Using the Comment feature available as an editing tool in Word, I inserted at least twenty comments, each addressing a different question to my client. After each comment, I inserted my proposed amendments to the contract, using Word's Track Changes feature.

When I was finished, I sent the documents, comments and changes included, to the client. The next day, after his review of my changes, he faxed me a document showing his responses and suggestions. I then revised the original form using Track Changes and sent it to him once again. In this way, after a series of exchanges, we had a final document which satisfied both of us. I sent that to him electronically. He called to say that he would take my final draft to the head of the hospital and discuss the changes with him. A few days later, I received from my client an executed document showing the hospital's execution of the document with our changes.

But what if I had sent the document to counsel for the hospital instead of to my client? And what if I had done so without reviewing my original comments and changes and "scrubbing" (deleting) them?

Under New York's view of metadata, the answer would appear to be that I would inadvertently have transmitted with my document metadata which would, in view of the extensive exchange of comments and proposed changes between me and my client, have revealed my client's confidences - especially his

extreme anxiety to secure employment immediately under almost any circumstances and his decision not to ask for changes that might jeopardize his chances of getting the job. I would appear specifically to have run afoul of NYSBA Opinion 782, *supra,* which concludes as follows:

> Lawyers have a duty under DR 4-101 to use reasonable care when transmitting documents by e-mail to prevent the disclosure of metadata containing client confidences or secrets.

I am not alone in skirting the risks imposed by metadata. In many cases, even those who have taken every precaution to "scrub" metadata from a document have instead been ensnared by its tentacles. In 2004, for example, U.S. military commanders in Iraq issued a report in Adobe PDF about the death of an Italian agent who was escorting a freed hostage through a security checkpoint. The report, heavily edited by the military for public release, was posted on an Army website. Within hours of its release, several readers had discovered that if the document was copied and pasted into Microsoft Word, all of the expunged elements would reappear. These included classified information about the procedures for manning and securing Army checkpoints.

Other cases of misdirected metadata abound. In 2004, the SCO Group sued DaimlerChrysler for breach of contract involving SCO's UNIX license. Metadata revealed that SCO's lawyers had originally intended to name the Bank of America as the defendant, but that the Bank had been dropped from SCO's complaint before the action was started. In another case, the Department of Justice was asked to remove from one of its web documents the name and social security number of a woman involved in an immigration case. They removed the social security number, but a PDF on the site continued to reveal the number when the PDF was converted to the "text version".

What's A Lawyer To do?

Fortunately, Microsoft, Corel, Adobe and others have recognized the need to cleanse or "scrub" an electronic document of metadata before it is sent, and they have supplied us with tools intended to accomplish that. Microsoft for example, offers an add-in which enables you to "permanently remove hidden and collaboration data, such as change tracking and comments, from Word 2003/XP, Excel 2003/XP, and PowerPoint 2003/XP files" (click on **www.tinyurl.com/6W8zl** to review and download the add-in). Many commentators have noted that documents in Microsoft Excel are an unusually extensive source of metadata. Microsoft Office Online offers a discussion of metadata and of the Microsoft add-in by Barbara Payne of the Payne Consulting Group. The article is headed, Control metadata in your legal documents (click on **www.tinyurl.com/27ngmm**).

For those lawyers who work in WordPerfect, Corel distributes instructions on the process of *"Saving WordPerfect Files Without Metadata,"* in an article by Laura Acklen. The process involves using the program's *edit>undo/redo history* feature (click on **www.tinyurl.com/y55pqq**). A company called MentorDocs (click on **www.tinyurl.com/272sto**) offers a free Document Assembler which contains macros for both Word and WordPerfect that enable the user to remove metadata without the need for a stand-alone removal program. The macros can be installed separately.

Metadata and PDF

Most users assume that conversion of a document to PDF destroys all metadata within the original document. This is not always the case and a prudent lawyer will want to guard against the possibility

that metadata lurks in the PDF. An Adobe Blog dated October 27, 2005 discusses the problem.

> For the most part, PDF is immune to these issues. PDFs represent the visual display as it will be printed. Still, it is a good idea to understand what if any risks are associated with PDF and metadata.

The blog was written before the release of Acrobat 8 which, according to Adobe, includes an "Examine Document" tool that offers powerful metadata removal. For those of you, who do not have Acrobat 8, click on **www.tinyurl.com/2rjem8** to view Adobe's suggestions for controlling metadata in PDFs. The National Security Agency (NSA) has recognized the problems inherent in the conversion of documents to PDF. In Report #1333-015R-2005 (12/13/2005), the Agency discussed the topic *Redacting with Confidence: How to Safely Publish Sanitized Reports Converted from Word to PDF.* In the Report, NSA summarizes the ways in which efforts to cleanse or scrub an article can fail; recommends complete deletion, not redaction, as the only effective scrubbing tool; and sets forth a detailed procedure for sanitizing a word document and copying it safely into PDF.

Recommended by the ABA

In a 41-screen slide show, the ABA has constructed a dramatic and instructive review of metadata and its inherent problems. *See, Metadata (and other things that go bump in the night),* by Catherine Sanders Reach, Director, American Bar Association Legal Technology Resource Center (July 27, 2006). After describing metadata, the slides encourage law firms to "get the facts, understand the implications, and establish firm policies to protect yourself and your clients." The slides then proceed through a series of illustrations showing how metadata becomes lodged in documents and offering access to ABA Ethics Opinions devoted to metadata.

One slide in the series lists and describes various software programs that are designed to work with Word, WordPerfect and PDF to eliminate metadata. The list includes all the following:

Payne's Metadata Assistant (**www.tinyurl.com/e2zef**) . The Assistant is designed to be used with Microsoft Word, Excel and PowerPoint. It comes in two versions, Retail and Enterprise (for multiple users). The program can analyze/clean files attached to outbound e-mail messages and convert them to PDF format for extra protection.

iScrub (Esquire Innovations, Inc.)
(**www.tinyurl.com/28e38p**) . iScrub is also designed for use with Microsoft products. iScrub removes the visible document properties and scrubs the more-difficult-to-reach file elements, such as the list of past authors (all document authors) and Deleted Text.

ezClean (Kraft, Kennedy & Lesser)
(**www.tinyurl.com/34asd2**) . This product is also limited for use with Microsoft products. Its website offers a review of metadata and some Microsoft bugs that have to be considered when using the program. It also gives detailed installation instructions.

Workshare Protect (Workshare) (**www.tinyurl.com/2ornnp**). This product offers real-time warnings when your outgoing e-mail contains passwords, social security and credit card numbers, or a history of deleted texts or comments. It removes metadata with the click of a button, or automatically with no interaction from the user.

Out-of-Sight (Softwise) (**http://tinyurl.com/2cn6mr**). Out-of-Sight offers compatibility with Office 2007 in its next release, and an evaluation license good for thirty days.

The ABA slide show also identifies Appligent as a source for metadata removal products. The company offers products which are designed to insure the integrity of PDFs. (Click on **http://tinyurl.com/3b8gf7** for a list of the company's products.

Lazar Emanuel is the Publisher of NYPRR. Research for this article was conducted and contributed by David Emanuel.

© Copyright 2008 −The New York Professional Responsibility Report (NYPRR)