Corporate & Commercial
Cite as: Jembaa Cole, *When Invisible Electronic Ink Leaves Red Faces: Tactical, Legal and Ethical Consequences of the Failure to Remove Metadata*, 1 Shidler J. L. Com. & Tech. 8 (Feb. 2, 2005), *at* <http://www.lctjournal.washington.edu/Vol1/a008Cole.html>

WHEN INVISIBLE ELECTRONIC INK LEAVES RED FACES: TACTICAL, LEGAL AND ETHICAL CONSEQUENCES OF THE FAILURE TO REMOVE METADATA

By Jembaa Cole[1]

ABSTRACT

There have been several instances in which seemingly innocuous metadata has wreaked professional and political havoc. Every electronic document has an invisible set of identifying data, its metadata. This Article explores consequences of ignoring metadata and suggests effective ways to tame metadata.

TABLE OF CONTENTS

INTRODUCTION

<1> In March 2004, attorneys for The SCO Group (SCO) were chagrined when a journalist revealed that he had been able to access earlier drafts of a complaint naming a different defendant when he examined electronic versions of SCO's copyright infringement complaints against DaimlerChrysler and AutoZone. The CNET News.com journalist reported:

> A Microsoft Word document of SCO's suit against DaimlerChrysler, seen by CNET News.com, originally identified Bank of America as the defendant instead of the automaker. This revision and others in the document can be seen through powerful but often forgotten features in Microsoft Word known as invisible electronic ink.[2]

<2> Many attorneys are not aware of existence of "metadata," the technical term for the "invisible electronic ink" uncovered by the journalist.[3] Such ignorance may lead to a similar public humiliation or worse if an attorney inadvertently discloses sensitive information

through metadata.

<3> The term metadata was unknown to the average person until several very public gaffes brought it to the forefront. In one of the most ironic incidents involving metadata, the Microsoft Company posted a downloadable copy of its 1999 Annual Report on its web site.[4] This report was a Microsoft Word document, and it contained metadata that showed that a portion of the document had been prepared on a Macintosh computer.[5] Microsoft has since posted tutorials about minimizing metadata in Microsoft Office Suite documents.[6]

<4> Earlier this year the Tony Blair administration also learned a painful lesson in metadata management.[7] The United Kingdom government placed a report about Iraq's weapons of mass destruction on its web site.[8] Contrary to the government's assertions, the document was drafted by civilians who had plagiarized the information from a student's thesis.[9] Further, the UK government asserted that the dossier was original and current; however, the metadata revealed that the document was merely a collection of documents written more than a decade earlier.

<5> Another incident occurred in the United States when The Washington Post inadvertently allowed private information from the DC sniper's demand letter to be revealed.[10] In his letter, the sniper detailed his failed attempts to contact the police. The letter listed the names and telephone numbers of every officer he had contacted. The Washington Post published this letter in PDF (portable document file) format document using black rectangles to redact the sensitive information. These rectangles were later removed and the full text of the letter was available on the Internet.

<6> These are a just a few of the worst case scenarios. There are many more ways metadata can trap the unwary.

WHAT IS METADATA?

<7> Metadata is quite simply data about data,[11] or hidden statistical information about a document that is generated by a software program. Software programs such as Microsoft Office Suite programs generate metadata on every document created.[12] Such information includes the document's author, date and time of creation, edits and editors, date and time the document was last accessed, deleted information, as well as previous versions of the document.[13] Although the average user does not see a document's metadata, it is always present and easily accessible. A document's entire history and background material will be included in its metadata.

<8> Metadata can be very useful. It provides a group of searchable data.[14] This is of utility, for example, in library sciences and web site

[15]

creation and management. Librarians use metadata to expedite searches. Searches are faster where a search engine targets a small number of key words and data, instead of the entire document. Web sites use metadata to lead search engines to their sites and potential sponsors use it to rate web sites.[16] The value of a web site is directly proportional to its accessibility.[17] It follows that the most valuable web sites are those with extensive metadata fields.

METADATA AND THE PRACTICE OF LAW

<9> Attorneys rely on electronic media to produce documents and to communicate; this increases the risk of inadvertent disclosure of sensitive information. Attorneys frequently prepare new documents and pleadings based on a template or earlier documents. These templates will retain all previous information as metadata. For example, the previous client's name and facts may be in the original document. If such a document were sent to a new client it might not only be embarrassing, it might even be a breach of the attorney's duty of confidentiality.[18]

<10> In complex contract negotiations, a number of attorneys will review one electronic contract. This contract may undergo numerous revisions from both sides of the negotiating table. Common revision tools generate long-lasting metadata. The "insert comment" function of Microsoft Word allows a reader to insert comments directly into a document.[19] Even if the comment is later deleted, it remains part of the document as unseen metadata. This information will be available to a third party who receives the document. Further, the "track changes" function will memorialize changes made to a document.[20] These changes may reveal a party's negotiation strategy and weaken its bargaining power. In more serious situations privileges may be jeopardized.

<11> In litigation, parties increasingly file, serve, or produce electronic copies of documents and pleadings.[21] Electronic discovery is less expensive and faster; however, there is a greater risk that a party will inadvertently provide sensitive or damaging information. Transmitting an electronic document to opposing counsel may compromise a case if the document's metadata contains sensitive information.[22] Metadata may also impact the issue of notice. For example, it would be embarrassing for a party to allege that they did not have notice of a particular fact when a document's history shows that they accessed a document containing that very information.[23] This party's date and time of access will remain part of the document's history.[24]

<12> These risks can be significantly minimized with a few simple changes. These changes should be a part of a larger document management policy. Implementing a policy to manage metadata is outlined below.

HOW TO EFFECTIVELY MANAGE METADATA

<13> The best way to manage metadata is to develop and implement an effective policy. This policy should be practical and should aim to create awareness of metadata and teach effective ways to manage or remove metadata. When developing a policy, consider the following:

<14> **Education**. This is the most important aspect of an effective policy. Introduce employees to metadata and explain its source, characteristics and utility. Such a presentation should give concrete examples of metadata in commonly-used documents. For example, demonstrate that a recycled document retains a previous client's information.

<15> Caution against the use of document processing short-cuts such as "fast-save" and "track changes." These features produce a large amount of long-lasting metadata. [25] Consider using third-party redline software such as Workshare® DeltaView®[26] instead of Microsoft Word's "track changes" feature. DeltaView allows comparison of redlined documents such as contracts.

<16> Consider including this step in your document management policy, new hire orientation and provide periodic reminders. Education and awareness is only the first step.

<17> **Minimize and Remove Metadata.** Computer software is programmed to produce metadata. Software options can be modified to reduce the amount of metadata created and to specify the fields of metadata that should be populated. These settings can be altered to minimize the amount of data produced. Microsoft publishes step-by-step guides to minimizing metadata fields in its software.[27] Although these guides are straight forward, some of the metadata fields are difficult to access;[28] as such this may not be the best option for effective removal of metadata. A more efficient alternative is to use a commercial metadata scrubber or document sanitizer.

<18> A metadata scrubber is a computer program that acts as a filter and strips metadata from documents. Documents can be filtered through these software programs before they are exported or stored.[29] As an added precaution some law firms use scrubbers that are set to automatically strip metadata whenever a document is saved. Keep in mind that the scrubber may not remove all metadata. A software developer will be able to design the most effective solution specifically tailored to your needs.

<19> As alternative to automated metadata scrubbers, consider converting shared documents into formats that do not support advanced metadata such as RTF (rich text format) or hard-copy. Note that these formats will still have some form of metadata, such as the name of the author. Although benign in most instances, there are situations in which this information is confidential. As such, a more effective method of metadata removal should be employed.

<20> Note that converting documents to PDF (portable document files) format does not eliminate the risk of inadvertent disclosure. For example, redacted text in PDFs is easily accessible.[30] It is as easy as changing the color of the text to white or cutting and pasting the text without the overlying black rectangle.[31]

<21> One can provide "clean" electronic documents without using metadata scrubbers by scanning.[32] Printing will eliminate metadata and the scanned version of the document will be clean.

<22> **Opt for Paper Copies.** The best way to remove metadata is to convert the document to a non-electronic format, paper copy. This is especially useful in discovery. Although electronic discovery has become increasingly popular, produce hard copies of responsive documents whenever possible. This will ensure that you are not inadvertently providing opposing counsel with sensitive information.

<23> **Use Clean Templates.** The policy should caution against recycling documents, rather using clean templates as the base of most documents. This will avoid the embarrassment of forwarding a client's information to another client. This has the added benefit of encouraging uniform firm documents.

<24> **Alternate Forms of Communication.** In many situations such as contract negotiations, alternate forms of communication are not practical, but wherever appropriate suggest alternate forms of communication such as paper copy, facsimile, telephone or mail.

<< Top

FOOTNOTES

1. Jembaa N. Cole, University of Washington School of Law, LL.M. Intellectual Property Law and Policy, Class of 2005. Thanks to Parag Gheewala, Jessica M. Neilson and Jane Winn for feedback on a draft of this article.

2. Stephen Shankland, *Hidden Text Shows SCO Prepped Lawsuit against BofA,* CNet News.com, *at* http://news.com.com/2100-7344-5170073.html (Mar. 4, 2004).

3. Matt Loney, *"Dodgy-dossier Syndrome" Rife in the Workplace*. ZDNet UK, *at* http://news.zdnet.co.uk/business/management/0,39020654,39117905,00.h (Nov. 14, 2003).

4. Scott Rosenberg, *Microsoft's Annual Report: Made on Macintosh,* Salon.com, *at* http://archive.salon.com/tech/log/1999/10/12/microsoft_report/ (Oct. 12, 1999).

5. *Id.*

6. How to Minimize Metadata in Microsoft Word Documents,

Microsoft Knowledge Base Article 237361, *at* http://support.microsoft.com/?kbid=237361 (last visited May 16, 2004).

7. Richard M. Smith, *Microsoft Word Bytes Tony Blair in the Butt*, *at* http://www.Computerbytesman.com/privacy/blair.htm (June 30, 2003).

8. Mark Ward, *The Hidden Dangers of Documents,* BBC News: Technology, *available at* http://news.bbc.co.uk/go/pr/fr/-/2/hi/technology/3154479.stm (last updated Aug. 18, 2003).

9. *Id.*

10. *Id.*

11. Unisa Web Authoring Guide, Using Metadata, *at* http://www.unisa.edu.au/wag/construct/metadata.asp (latest content revision Jan. 28, 2004).

12. William Robertson, *Protecting against Metadata Mishaps*, 21 No. 5 LJN's Legal Tech. Newsl. 1 (Aug. 9, 2003).

13. Jeffrey Beard, *Make Information More Difficult to Access,* 18 No. 3 Tex. Law. 19 (March 25, 2002).

14. Chris Taylor, An Introduction to Metadata, The Univ. of Queensland, Cybrary, *available at* http://www.library.uq.edu.au/papers/ctmeta4.html (Revised July 29, 2003).

15. *Id.*

16. Using Metadata, *supra* note 10.

17. Search Engine Marketing, Web Marketing Info Page, *available at* http://www.wilsonweb.com/webmarket/searchengine.htm

18. Wash. State Rules of Prof'l Conduct R. 1.6., *available at* http://www.courts.wa.gov/court_rules/?fa=court_rules.display&group=ga&set=RPC&ruleid=garpc1.06 .

19. Robert E. Bershad and Laura Bandrowsky, *Avoiding Microsoft Word Document Land Mines,* 20 No. 2 LJN's Legal Tech. Newsl. 1 (May 2002).

20. Mark Grossman, *Metadata Traps for the Unwary,* The Miami Herald, April 1, 2002, *available at* http://www.mgrossmanlaw.com/articles/mhtl/metadata.htm

21. David W. Snyder, *Report on Legal Technology: The Perils of Sharing: Cleaning Up Your E-Mail,* 24 pa. Law. 31,

31Sept./Oct. 2002, *available at* http://www.klettrooney.com/newsroom/attorney_articles/0208_snyder.pdf

[22] *Id.,* at 31.

[23] Grace V. Bacon, *Legal Analysis: The Fundamentals of Electronic Discovery,* 47 B.B.J. 18, 19 (2003).

[24] *Id*.

[25] Mark Grossman, *Metadata Traps for the Unwary,* The Miami Herald, April 1, 2002, *supra* note 21; Mark Ward, *The Hidden Dangers of Documents,* BBC News: Technology, (published August 18, 2003) *at* http://news.bbc.co.uk/go/pr/fr/-/2/hi/technology/3154479.stm, *supra* note 7.

[26] Workshare® DeltaView®, DMS Integration, *at* http://www.workshare.net/products/pr_dv_dms.htm

[27] How to Minimize Metadata in Microsoft Word Documents, Microsoft Knowledge Base Article 290945 (for Microsoft Word 2002), *at* http://support.microsoft.com/default.aspx?scid=kb;EN-US;290945; Microsoft Knowledge Base Article 237361 (for Microsoft Word 2000), at http://support.microsoft.com/?kbid=237361; Microsoft Knowledge Base Article 223790 (for Microsoft Word 1997) http://support.microsoft.com/default.aspx?scid=kb;EN-US;223790.

[28] *Id*.

[29] Eric Doyle, *Metawall Strips Hidden Word Data,* Computer Weekly, Dec. 12, 2002 at 30; Storm Evans, *Technology in Practice,* Product Watch, 29 Law Prac. Mgmt. 20 (May 2003/ June 2003) (Examples of metadata scrubbers are: Workshare® Synergy™, Metawall).

[30] David Hricik and Bob Jueneman, *The Transmission and Receipt of Invisible Confidential Information (2003)* http://www.hricik.com/eethics/Metadata1103.doc.

[31] *Id*.

[32] *Id*.

<< Top