RETURNING TO PLATO'S CAVE: Metadata's Shadows in the Courtroom

Zachary Rosenberg*

ABSTRACT

The computer revolution changed every facet of our lives, including litigation. Though computer interfaces are designed to display information through familiar renderings of everyday physical objects, computer files are stored and behave differently from their physical counterparts. Metadata, the information contained in a computer file that are almost invisible to users, can profoundly affect the admissibility and authenticity of digital files. This paper explains what metadata are and the role they play in litigation to authenticate other evidence or as evidence in itself. This paper proposes a new best practice for attorneys: whenever a lawyer receives an electronic file or hard drive from a client, the attorney should immediately back up and forensically image it so that the metadata are preserved.

[*] J.D. Candidate, 2016. Special thanks to Professor Bob Dauber, Clinical Professor at Arizona State University Sandra Day O'Connor College of Law, for all of his help and advice.

INTRODUCTION

The term "electronic information" invokes images of icon-littered desktops, "My Documents" folders, or perhaps the iconic W, X, P, and O of Microsoft's ubiquitous suite of applications we use, love, and love to hate.[1] Very few of us, certainly no one born after 1990, think of computer files like Owen Wilson did in the movie Zoolander. "The files are in the computer,"[2] his character, Hansel, proclaimed before throwing an iMac over a balcony expecting the "files" to fly everywhere when the computer hit the ground.[3] While most lawyers understand that computer files are not stored in hard copy inside their computers, they may not fully appreciate how different computer files are from physical files or how ignorance of those differences can impact litigation.

Electronically Stored Information (ESI) is any file or evidence stored in a computer or other electronic device.[4] Failing to preserve ESI is the conduct

[1] *E.g.*, Gregg Keizer, *Office 2016 for Mac Users Plagued by Crashes After Upgrading to OS X El Capitan*, PCWORLD (Oct. 2, 2015, 8:29 AM), http://www.pcworld.com/article/2988941/software-productivity/office-2016-for-mac-users-plagued-by-crashes-after-upgrading-to-os-x-el-capitan.html.

[2] ZOOLANDER (Paramount Pictures 2001).

[3] *See* Desert Mobile Service, *Computers in Film - How NOT to Get that File Out of the Computer - Two Morons Try to Retrieve Files*, YOUTUBE (Feb. 6, 2013), http://www.youtube.com/watch?v=sqVCS4gaJ5M.

[4] ESI "includ[es] writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations—stored in any medium from which information can be obtained either directly or, if necessary, after translation by the responding party into a reasonably usable form." FED. R. CIV. P. 34(a)(1)(A).

most likely to result in sanctions during e-discovery.[5] When lawyers or their clients fail to appreciate the important differences between electronic and hard copy versions of a file, even routine disputes can become fights over bad faith in the litigation. For example, in *Raines v. College of Greater Cleveland Inc.*, the plaintiff sued her former employer on contractual and discriminatory grounds after the university fired her.[6] In an amended complaint, the plaintiff alleged tortious spoliation of evidence[7] after the defendants deleted a report the plaintiff prepared while working for the defendant.[8] Even though the plaintiff had a hard copy of the report, she wanted the metadata on the grounds that they had independent significance.[9] The court denied the defendant's motion for summary judgment and opted to answer questions about the duty to preserve metadata at a later time.[10] This means that the trial court was open to hearing arguments about not just the relevance of metadata but the duty to preserve them. This case is significant because the discovery dispute was about information that is only stored in digital versions of documents. The court's ruling implies that there are, or at least could be, legally significant differences between a hard copy of a report and its electronic counterpart.

In *Long Bay Mgmt. Co. v. Haese LLC*, the Massachusetts Court of Appeals upheld the trial court's issuance of a default judgment against a law firm for failing to comply with a court order to produce the metadata of a billing file.[11] The metadata were needed to prove which alterations were made, and when, to the file.[12] In upholding the sanctions, the court impliedly rejected the defendant's argument that the defendants were not able to extract the metadata from the files at issue.[13] The court noted that the firm produced files

[5] Dan H. Willoughby, Jr. et al., *Sanctions for E-Discovery Violations: By the Numbers*, 60 DUKE L.J. 789, 803 (2010).

[6] Raines v. Coll. Now Greater Cleveland, Inc., No. 1:14-CV-00003, 2014 WL 2506261, at *1 (N.D. Ohio June 3, 2014).

[7] In Ohio, tortious spoliation of evidence requires that the defendant have knowledge of "(1) pending or probable litigation involving the plaintiff, (2) knowledge on the part of defendant that litigation exists or is probable, (3) willful destruction of evidence by defendant designed to disrupt the plaintiff's case, (4) disruption of the plaintiff's case, and (5) damages proximately caused by the defendant's acts." Smith v. Howard Johnson Co., 615 N.E.2d 1037, 1038 (Ohio 1993). Though this tort is not recognized by many jurisdictions, the same conduct could be sanctionable under the applicable rules of civil procedure.

[8] *Raines*, 2014 WL 2506261, at *1.

[9] *Id.* at *4.

[10] *Id.* at *5.

[11] Long Bay Mgmt. Co., v. Haese, LLC, No. 14–P–991, 2015 WL 7213811, at *3 (Mass. App. Ct. Nov. 17, 2015).

[12] *Id.*

[13] *Id.*

allegedly containing the metadata, that were in fact just copies of the final billing statement that were relabled as the old documents; the file creation date was four years after the timeframe in question.[14] This case shows how important it is for lawyers to understand what metadata are and what steps should be taken to preserve them.

Metadata have two important functions to play in litigation. First, metadata can authenticate or challenge the authenticity of other ESI. Even minute changes to the metadata or files themselves can irrevocably alter the document and potentially render it inadmissible. Second, metadata can be primary evidence, like the file in *Haese*, to show when, where, or how something occurred. Metadata are records of, *inter alia*, when files were created, modified, opened, or transmitted, and often by whom. All of which could be important evidence in a case.

Court-imposed sanctions for failing to preserve ESI increased sevenfold between 2003 and 2009.[15] Although only a handful of cases resulted in sanctions against counsel,[16] the escalation of awards against lawyers based on counsel misconduct for failing to preserve ESI poses a problem for lawyers and clients. In particular, handling electronic information the same way as physical documents runs a substantial risk of altering those files enough to potentially preclude their authentication, and thereby admissibility, in court.

Part I of this paper discusses ESI and metadata generally: what they are, how they work, how they can be altered, and why they may be relevant in litigation. Part II discusses legal frameworks for ESI and how they implicate metadata. It addresses the Federal Rules of Civil Procedure and the Federal Rules of Evidence and how metadata play, or can play, a role in the litigation process. It then discusses issues with confidentiality and waiver. Part III proposes a new best practice for lawyers dealing with clients' electronic files. Part IV concludes.

I. BACKGROUND

In order to understand the unique complexities of computer files, it is necessary to understand some of the illusions, or more accurately, the

[14] *Id.* at n.9.

[15] Willoughby et al., *supra* note 5, at 816. For example, the defendant in *Prezio Health, Inc. v. Schenk* was sanctioned after three e-mails were deleted. Prezio Health, Inc. v. Schenk, 3:13 CV 1463 (WWE), 2016 WL 111406, at *1 (D. Conn. Jan. 11, 2016). The court specifically reprimanded the defendant and his counsel for not advising the defendant's wife to take her new iPad to the Apple Store to be set up by an expert rather than attempting to do it herself. *Id.* at *3.

[16] *Id.* at 815.

misleading representations of information that computers present on screen.[17] Long before computers had user-friendly, colorful interfaces, computers used punch cards to store information.[18] Programming a computer involved selectively punching holes in these cards.[19] Computers eventually migrated to command line interfaces where a user would enter instructions on a keyboard to access applications and run programs.[20] However, that all changed in the 1980s after the invention of the graphical user interface (GUI, pronounced "gooey").

In 1984, Apple introduced the Macintosh,[21] and computer interfaces became skeuomorphic. A skeuomorphic computer interface is designed to emulate the physical world.[22] For example, the "desktop" looks like the top of a desk, replete with scattered files, folders, and a trashcan; word processors mimic typewriters by displaying what you type on a representation of 8 ½ X 11 inch paper; and e-mail programs still refer to sending a copy of an e-mail to a third party as a carbon copy ("CC")—an allusion to using carbon paper placed between two sheets of paper to "copy" the writing on the top sheet to the second sheet as the writer's pen pressed to the top sheet.[23] Skeuomorphism is helpful because it gives users a reference to understand what a computer does and how computer files are related to their physical counterparts.[24]

[17] For a more complete, though somewhat dated, explanation of computer terminology and forensics, see Craig Ball, *Beyond Data About Data: The Litigator's Guide to Metadata*, CRAIGBALL.COM (2011), http://www.craigball.com/metadataguide2011.pdf.

[18] Victor Fay-Wolfe, *History of Computers*, THE UNIV. OF R.I., http://homepage.cs.uri.edu/faculty/wolfe/book/Readings/Reading03.htm (last visited May 2, 2016).

[19] *See id.*

[20] MS-DOS is the classic example. Microsoft used it as the basis for its Windows line of operating systems.

[21] Xerox actually developed the first viable GUI in 1974, but it was never a commercial product. Michael Tuck, *The Real History of the GUI*, SITEPOINT (Aug. 13, 2001), http://www.sitepoint.com/real-history-gui/5/; *see also* Antisubliminal, *Apple-1984*, YOUTUBE (June 19, 2006) https://www.youtube.com/watch?v=R706isyDrqI.

[22] *Skeuomorphism*, TECHNOPEDIA, http://www.techopedia.com/definition/28955/skeuomorphism (last visited May 2, 2016). "Skeuomorphism is 'an ornament or design *representing a utensil or implement*.'" Thomas Q. Brady, *What is the Opposite of Skeuomorphism?*, QUORA (Jan. 23, 2013), http://www.quora.com/What-is-the-opposite-of-skeuomorphism.

[23] *See Carbon Copy*, DICTIONARY.COM, http://dictionary.reference.com/browse/carbon--copy (last visited May 2, 2016).

[24] However, the last few years have seen computer interfaces begin to turn away from these design elements and embrace alternative design languages. For example, recall the massive change to Apple's mobile software between versions 6 and 7 in 2013. The massive change in the graphic schemes represented a shift away from many of the gimmicky skeuomorphic graphics elements to "flatter" more dynamic ones that had much less grounding in real world objects.

Despite the friendly interface and comfort of familiar representations of computer functions, skeuomorphism obscures the nature of computer files by giving the impression that they look and behave like physical files. The final content we see, read, and edit is the most important part of the documents we use; however, there are parts of computerized documents that do not behave like their real world counterparts—when such counterparts even exist.

Though computers display documents and information as though they were part of a physical world, the files themselves are entirely different and far more dynamic than paper files. Consider Plato's cave, a thought experiment in the seventh book of *The Republic*.[25] Plato's thought experiment began with a group of prisoners who lived their entire lives chained in a cave, facing a wall.[26] A fire burned behind them, casting their shadows upon the wall.[27] Since their chains prevented them from turning to see the light, they only knew shadows and mistook them for the real objects.[28] One of these prisoners escaped and entered the real world.[29] His perception of reality was completely changed by his exposure to sunlight and the world beyond the cave; he understood that the shadows were not objects or people but rather distorted representations of real objects.[30] When he was recaptured and returned to the cave, he discovered he could no longer perceive the shadows as he once did—and how his fellow prisoners still did—nor could his fellow prisoners understand his perspective.[31] The same could be said of computer information: what we see on screen is like a shadow on the cave wall; we cannot see or perhaps even appreciate that they are obscured representations of the digital world in our devices, projected onto our screens, not by fire but by sophisticated software and specialized hardware.[32]

ESI is more than just the content of documents; it includes detailed information about when files were created, by whom, where, by what software, on what computer—in some cases, previous versions of the document may even be contained in a single file, but completely invisible

[25] PLATO, THE REPUBLIC BOOK VII, *reprinted in* THE COLLECTED WORKS OF PLATO 747, 747–52 (Huntington & Cairns eds., 1980), http://www.anselm.edu/homepage/dbanach/repub7.htm.

[26] *Id.*

[27] *Id.*

[28] *Id.*

[29] *Id.*

[30] *Id.*

[31] *Id.*

[32] Ralph Losey, *Plato's Cave: Why Most Lawyers Love Paper and Hate e-Discovery and what This Means for the Future of Legal Education*, E-DISCOVERY TEAM (Aug. 8, 2009), http://e-discoveryteam.com/2009/08/11/platos-cave-why-most-lawyers-love-paper-and-hate-e-discovery-and-what-this-means-to-the-future-of-legal-education/.

unless the user looks for them. Some files, pictures in particular but even some note taking and word processing applications, include GPS or other location data to identify where a photo was taken or where a document was created.[33] Though this information is hidden to the user, some of this data can be easily and even inadvertently modified, raising potential questions of spoliation and authentication.

A. What Is Electronically Stored Information (ESI)?

As the name suggests, ESI is any information or digital data,[34] including e-mails, documents, webpages, databases, and images stored on a tablet, smartphone, sever, or computer or one of its peripheral devices (hard drives, CDs, DVDs, flash drives, etc.).[35] In the last 40 years, physical files have been largely abandoned for computer files. Between 1986 and 2007, global computer storage capacity grew an average of twenty-three percent per year.[36] In 2002, for the first time in history, more information was stored digitally than physically,[37] and by 2007, about ninety-four percent of all information in the world was stored electronically.[38] To provide some context, this means that the amount of digital information in the world has grown at five times the rate of global Gross Domestic Product.[39] Every minute of every day, people send 200 million emails, upload forty-eight hours of video to YouTube, run 2 million Google searches, and create 571 thousand new websites.[40] Though the best available, even these statistics date back to 2007

[33] *E.g.*, *How to See the Exact Location Where a Photo was Taken with a Mac*, OSXDAILY (May 8, 2015), http://osxdaily.com/2015/05/08/view-exact-location-photo-taken-preview-mac.

[34] The Federal Rules of Civil Procedure define ESI and physical evidence as "including writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations—stored in any medium from which information can be obtained either directly or, if necessary, after translation by the responding party into a reasonably usable form." FED. R. CIV. P. 34(a)(1)(A).

[35] BARBARA J. ROTHSTEIN ET AL., FED. JUDICIAL CTR., MANAGING DISCOVERY OF ELECTRONIC INFORMATION: A POCKET GUIDE FOR JUDGES 2–4 (2007), http://www.fjc.gov/public/pdf.nsf/lookup/eldscpkt.pdf/$file/eldscpkt.pdf.

[36] Martin Hilbert, *How Much Information Is There in the World*, SCIENCE DAILY (Feb. 11, 2011), http://www.sciencedaily.com/releases/2011/02/110210141219.htm.

[37] *Id.*

[38] *Id.*

[39] Martin Hilbert, *How Much Information Is There in the "Information Society"?*, 9 SIGNIFICANCE, no. 4, 2012, at 9, http://onlinelibrary.wiley.com/doi/10.1111/j.1740-9713.2012.00584.x/abstract.

[40] Mitch Monsen, *The Explosive Growth of Digital Data*, TOP TEN REVIEWS (Feb. 25, 2013), http://ftp-hosting-services-review.toptenreviews.com/the-explosive-growth-of-digital-data.html.

when Apple had just released the first iPhone; the smartphone revolution had barely begun, tablets were at most a niche market, and wearables[41] were barely a glint in engineers' eyes. The amount of computer information in the world now doubles about every eighteen months; thus, the numbers are much higher now.[42]

While it is easy to think of digital information as analogous to its physical counterparts, digital information is completely different in how it is stored, processed, and organized. One of the most significant differences is metadata.

B. What Are Metadata?

There are many misconceptions about what information are metadata. Metadata are generally described as data that describe data, or data about data.[43] The term metadata properly refers to information about a file that is not the content of the document.[44] For example, when you create a new word

[41] Wearables are devices that the user wears or implants somewhere in the body; they include devices like the Apple Watch, FitBit, Nike+, Heart Monitors, etc. Often these devices rely on other devices like smart phones to store information and track trends in the data they create. Depending on the device, this information may be stored on the internet (in the "cloud") or even shared over social networks. *See generally* Kiana Tehrani & Andrew Michael, *Wearable Technology and Wearable Devices: Everything You Need to Know*, WEARABLE DEVICES MAG., http://www.wearabledevices.com/what-is-a-wearable-device (last updated Mar. 26, 2014).

[42] Jeff Vance, *Big Data Analytics Overview*, DATAMATION (June 25, 2013), http://www.datamation.com/applications/big-data-analytics-overview.html.

[43] Arlen L. Tanner, *Metadata: Why the Fuss? A White Paper on Metadata*, BLOOMBERG LAW REPORTS (2011), http://www.shb.com/~/media/files/professionals/tannerarlen/metadatawhythefuss. One legal textbook defined it as "information stored by applications such as word processing programs and spreadsheets that enhances the functionality of the software in ways that are invisible to the user." GEOFFEY C. HAZARD, JR. ET AL., THE LAW AND ETHICS OF LAWYERING 380 (5th ed., 2010). In the Maryland District Court's suggested protocol for handling ESI:

> "Meta-Data" means: (i) information embedded in a Native File that is not ordinarily viewable or printable from the application that generated, edited, or modified such Native File; and (ii) information generated automatically by the operation of a computer or other information technology system when a Native File is created, modified, transmitted, deleted or otherwise manipulated by a user of such system. Meta-Data is a subset of ESI.

Suggested Protocol for Discovery of Electronically Stored Information ("ESI"), U.S. DISTRICT CT. FOR THE DISTRICT OF MD. 2–3, http://www.mdd.uscourts.gov/news/news/ESIProtocol.pdf (last visited May 15, 2016).

[44] *See generally* SHIRA A. SCHEINDLIN ET AL., ELECTRONIC DISCOVERY AND DIGITAL EVIDENCE: CASES AND MATERIALS 249–66 (2009). Metadata are "[i]nformation about a particular data set or document which describes how, when, and by whom the data set or document was collected, created, accessed, or modified; its size; and how it is formatted. Some metadata, such as file dates and sizes, can easily be seen by users; other metadata can be hidden from users but

processing document, you save it to the computer. Apart from the text you have written and all of the formatting information (margin size, type-face, font size, font color, etc.), the file contains information like the creation date and time, location in the computer, the physical location of the computer that created it, keywords or tags, when it was last modified, who created it, who can access it, and more.[45] For the most part, the average user does not need to access or view metadata, but they can be useful. For example, if there are multiple versions of a report, you may look at the date a file was modified or created to find the latest version or whichever draft you are seeking. While metadata generally are important for computers to function, developers to troubleshoot errors, and IT personnel to solve problems, most metadata are not useful to litigators.[46]

Though lawyers often use the term "metadata" broadly to encompass any information that is not visible to the user when a file is open, most of the information that lawyers fret over is technically plain data:[47] tracked changes,

are still available to the operating system or the program used to process the data set or document." ROTHSTEIN ET AL., *supra* note 35, at 24–25.

[45] Tanner, *supra* note 43.

[46] *Id.* Metadata includes information about the size and dimensions of a photo, the encoding scheme used to compress a file, and other information that enables a computer to interpolate the data within a file so it can be properly displayed. This is how computers "know" which software program to use to open a file.

[47] User-generated information is properly referred to as data, not metadata. *Id.* A common example of this invisible information is tracked changes. By hiding the various changes made by the reviewer(s), *id.*, the data are still part of the file but not immediately visible. Microsoft refers to this as viewing the "final" version. Tracked changes can be set to display all, some, or none of the changes entered by editors. Tracked changes is still on but set to display in the "final" view and not the "markup" view; the information is saved in the file but not visible to the user. Another example is hiding a column or row in a spreadsheet. Just because it is not visible does not transform the information into metadata; it is still part of the user-created content of a document. Likewise, redacting a document by "highlighting" a word or sentence in black may hide the information to the user at first glance, but the redacted text is still part of the file. Metadata, in contrast, are usually just as invisible to the viewer as hidden information, but are not directly created or changed to reflect the information inside the document. However, the metadata may change to reflect changes in the document's size, location, modification, etc.

Simple as this sounds, people often mistakenly hide information believing it is deleted. For example, the Transportation Security Administration (TSA), perhaps attempting to analogize how their full body "naked" scanners work, inadvertently released a "redacted" document to the public that included the full text of the redacted words; to view the full text, one only had to copy and paste the text into a word document. William Deutsch, *How to Redact a PDF File*, ABOUT MONEY, http://bizsecurity.about.com/od/informationsecurity/ht/How-To-Redact-A-Pdf-File.htm (last updated Dec. 10, 2014); *see also* Ward Room Staff, *Redactions Revealed: The Six Secrets You Need to Know from the Obama Subpoena Request*, NBC CHICAGO (Apr. 22, 2010, 3:25 PM), http://www.nbcchicago.com/blogs/ward-room/The-Six-Secrets-You-Need-to-Know-From-the-Blagojevich-Filing-91848634.html ("redacting" sensitive information from a subpoena of President Barack Obama by using the black highlight feature of Microsoft word lead to the hidden rows or columns, redacted information, past versions of the document, etc.[48]

There is no direct analogy for metadata in the physical world, but they are closest to the Dewey Decimal System in libraries. If the book's content is the data, changing, deleting, or redacting any part of it is different from changing the numbers printed on the spine, its location in the library, or the list of people who have checked out the book.

This analogy is imperfect, however, because, unlike the Dewey Decimal System, metadata can be altered by interactions that would not change the decimals of a book. Opening the book or penciling a note in the margins (though a violation of library rules) does not change the Dewey Decimal; such behavior, though slight or even unintentional, changes the metadata of a computer file.

Occasionally, lawyers may use the term "metadata" to refer to files that have been deleted by the user but still exist in the "empty" space on a computer. Unless you set up your system otherwise, deleting a file by "emptying the trash" does not actually delete the file. The pointer files or reference data, a specific type of metadata, are removed. The file still exists in the "empty space" on your hard drive, but the computer has removed all reference to it—a kind of compartmentalization. The "deleted" file still becomes a proverbial "elephant in the room"; the computer acts like the file does not exist even though it is still right where it was before it was deleted.

Deleted files remain in this empty space until they are overwritten by new information.[49] Even though this information is not visible, it is still a mix of data and metadata; deleting files does not turn their ghosts into metadata; it just hides the document's data. Differently stated, deleting a computer file deletes key pieces of metadata describing where the file is located, what it is called, and where it is stored on the hard drive.[50]

embarrassing public disclosure of sensitive information); *The Censored Elements of the Report on the Death of Nicola Calipari*, VOLTAIRE NETWORK (June 8, 2005), http://www.voltairenet.org/article30249.html (discussing how the Department of Defense made the same error and failed to redact sensitive information about the death of Nicola Calipari, the head of Italian Secret Service); *c.f.*, ARCHITECTURES & APPLICATIONS DIV. OF THE SYS. & NETWORK ATTACK CTR., NSA, REDACTING WITH CONFIDENCE: HOW TO SAFELY PUBLISH SANITIZED REPORTS CONVERTED FROM WORD TO PDF 4–14 (2008), http://www.nsa.gov/ia/_files/support/I733-028R-2008.pdf (NSA discussing how to redact sensitive information).

[48] Tanner, *supra* note 43.

[49] As you fill up your computer, these files are partially or completely overwritten. Moving files, defragmenting your hard drive, and other routine functions can overwrite part or all of these deleted files.

[50] As a real world example, consider the last time your computer shut down while you were working on an unsaved document. You may have used a file recovery program to search for

This distinction between data and metadata is important because the information commonly at issue in confidentiality questions is data, not metadata.[51] The invisible, user-generated information in a file is not metadata and can at most be called "pseudo-metadata."[52] Metadata are generally not created or entered by the user, though users can modify them if they wish.[53] There are two key types of metadata: metadata stored in the file and metadata stored in a separate file.[54]

1. Metadata in Separate Files

Metadata contained in separate files are often most relevant for file management systems or databases.[55] This metadata allow for large volumes of files to be searched by keyword, author, date created, or other criteria.[56] Search engines and some computer search features maintain index files that include information about many (or even all) of the files in a computer. Metadata of this fashion will not, generally, be transmitted to lawyers unless it is specifically at issue; however, some file information, including the name of the document, is stored in a separate file from the document itself.

For example, if I saved the draft of this article in the "law review" file on my laptop; the file itself is not stored in a folder called "law review"; it is stored on the computer by an entirely different method based on how my hard drive is formatted. In fact, the file may not be stored as one file; it may be spread out into many different fragments throughout the hard drive.[57] The metadata translate the organization scheme of the hard drive into the

the document. If you found that document, it may have been part of this ghost data in the empty part of your hard drive. This metaphor is imperfect because failing to save a file usually means that a copy was not transferred from the RAM to the Hard Drive. However, some software programs do automatically save copies to the hard drive at regular intervals. *See infra* note 59 and accompanying text.

[51] Tanner, *supra* note 43.

[52] *Id.* Some have referred to this as substantive metadata. W. Lawrence Wescott II, *The Increasing Importance of Metadata in Electronic Discovery*, 14 RICH. J.L. & TECH., no. 3, 2008, at 3.

[53] Backdating a file is the user modifying, or attempting to modify, the metadata. There are also programs that allow files' metadata to be modified, viewed, or changed in batches; it is most common for pictures and images. *See, e.g.*, *DicomBrowser*, NEUROINFOGRAPHICS RES. GRP., http://nrg.wustl.edu/software/dicom-browser (last visited May 2, 2016).

[54] Tanner, *supra* note 43.

[55] *Id.*

[56] *Id.*

[57] When computer files become too broken up or disjointed, it can slow down computer functions; the computer has to work harder to assemble the complete file to display to the user. This is called fragmentation or a fragmented drive.

organization scheme graphically presented on screen.[58] A metadata file states what the file is named, where the file is located within my customized folders, and where the file is written on the physical hard drive. When you “open” a document, it is copied off of the hard drive and combined from the disconnected pieces into a single, unified file in the Random Access Memory (RAM) of a computer. When you work on a document but have not saved it, you are working off of this copy stored in the RAM of a computer. Unlike hard drives, RAM is erased whenever the electric current running to it is interrupted. If the computer crashes or loses power, the RAM is erased and the data is lost. When you “save” the document, the version stored in RAM is copied over the version saved on the hard drive.[59]

2. Metadata in Files

There are two types of metadata contained within a file: file system metadata and program metadata. File system metadata (system metadata) are metadata inserted by a computer’s operating system to aid in handling the file.[60] System metadata include information like the date modified, date created, date accessed, date printed, location in the system, location on the hard drive, and other information about the file that enables the computer’s operating system to properly open and store it.[61] Some of this information is more reliable than the rest, as system settings may cause these data to change based on different user behaviors.[62] System metadata are generally consistent across all file types on a given computer because they are generated by the computer’s operating system and not individual applications.[63]

Program metadata[64] are information inserted into the file by the application that created, opened, or modified the file.[65] Thus, the metadata in files will

[58] The folders in your computer are themselves an illusion of how files are stored on a computer. Files often are not even stored in one cohesive block within a computer’s hard drive; they can be stored in pieces on the disk. The fuller your hard drive, the more likely that files will be broken up into more pieces to fit in the free space. When files are broke up into too many pieces, the computer can become slow. This is called fragmenting.

[59] The act of opening the file will change the metadata indicating when it was lasted opened or accessed, and saving the file will change the “modified date.” By contrast, when you “delete” a file from your computer, this reference file is only information that is actually deleted. The original file remains in its block(s) spread throughout the hard drive, but those blocks have been reclassified as “empty” by the computer.

[60] Tanner, *supra* note 43.

[61] *Id.*

[62] *Id.*

[63] *See id.*

[64] Sometimes referred to as OLE Metadata.

[65] Tanner, *supra* note 43.

vary from program to program and even file to file, especially if multiple programs open the same file. For example, prior to Microsoft Office 2003 (version 11), Microsoft Word retained a list of all previous authors in every Word document.[66] When opening an old document in a newer version, that metadata is deleted. This is just one example; computer programs can be written to insert whatever metadata the developer wants; thus, it is impossible to create a comprehensive list of metadata fields.

3. Changing Metadata

Metadata is not foolproof and can be intentionally or inadvertently modified.[67] Changing the computer's date and time then opening or modifying a file will change the last modified date to whatever the clock has been reset to. However, this does not mean all the metadata are changed to reflect the changed date and time. There may be inconsistencies. An expert can often identify when someone has tried to manipulate metadata, but someone with enough expertise could conceivably create a perfect fabrication. Though not tamper proof, metadata are often more reliable than conventional authentication methods.

"Imaging" a file by exporting it into a TIFF[68] or PDF[69] will usually remove invisible data from the image file, depending on the settings.[70] These types of "static format"[71] documents may be sufficient for some lawsuits, depending on the dispute, and where appropriate, these documents could be augmented by a printout of specific metadata.[72] Producing a document in one of these

[66] *Id.*

[67] Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 548 (D. Md. 2007) (citing to other sources).

[68] TIFF is a standardized file format for storing scanned images. Originally created in the 1980s and largely unchanged since 1992, the standard is controlled by Adobe. *TIFF Versus PDF – An Overview of Their Merits*, AQUAFOREST, http://www.aquaforest.com/en/tiff_versus_pdf.asp (last visited May 2, 2016).

[69] PDF is an open standard originally created for file transfer between computers and programs. *Id.*

[70] *Id.*

[71] Static format documents are "fixed" and will not change. *See generally* Margaret Rouse, *Dynamic and Static*, TECHTARGET, http://searchnetworking.techtarget.com/definition/dynamic-and-static (last visited May 2, 2016).

[72] *E.g.*, CQuest Am., Inc. v. Yahasoft, Inc., No. 13-cv-3349, 2015 WL 4576778, at *3, *6 (C.D. Ill. July 30, 2015) (reviewing a motion for sanctions when Yahasoft provided a PDF of the source code without the associated metadata and declining to provide harsher sanctions but requiring the production of the information in its native format); *see also* Kathy Perkins & Dave Deppe, *"Byte" Me! Protecting Your Backside in an Electronic Discovery World (Not Just for Litigators)*, 76 J. KAN. B. ASS'N 22, 31 (2007).

image formats is often called "scrubbing" a document. It removes hidden data (tracked changes and hidden fields), and all the metadata (date created, file name, file location, date modified). Usually when scrubbing a file, a lawyer is most concerned about protecting confidentiality, but imaging the files scrubs more than just privileged content. By redacting a file, the content of the file has been sufficiently modified to raise potential authenticity questions. One under-appreciated facet of computer files and metadata is how susceptible they are to inadvertent modification in ways that can hamper proper authentication and identification of files.

4. Hashing

A specific type of metadata, known as hashes (also called hash values or hash functions), can be particularly helpful in litigation—but also can be easily modified. Hashes are computer file equivalents of human fingerprints; they uniquely identify a computer file but do not actually say anything about a file's contents.[73] Just as having someone's fingerprint does not tell you anything about the person or give you enough information to make a clone, file hashes allow you to identify unique computer files without having to manually compare the files.[74] A hash value uniquely identifies a file, group of files, or portion of a file using a mathematical algorithm.[75]

> The most commonly used algorithms . . . will generate numerical values so distinctive that the chance that any two data sets will have the same hash value, no matter how similar they appear, is less than one in one billion.[76] "Hashing" is used to guarantee the authenticity of an original data set . . . .[77]

To understand the importance of hash values, consider how we distinguish between different people. If you have someone's fingerprint or DNA, you can uniquely identify that individual; however, you need more information to

[73] Craig Ball, *In Praise of Hash*, *in* MUSINGS ON ELECTRONIC DISCOVERY 56, 56 (2013), http://www.craigball.com/BIYC.pdf.

[74] *Id.* at 56–57.

[75] ROTHSTEIN ET AL., *supra* note 35, at 24. At a technical level, hashes, or hash functions, are algorithms that translate plain text information into a sequence of numbers. Indexes and computer search features use these hashes to quickly locate information by translating the plain text the user enters in a search box and comparing it to the hashes for all of the files in the computer's or database's index. Databases will assign each file a unique hash key. *What Are Hashes*, WISEGEEK, http://www.wisegeek.com/what-are-hashes.htm (last visited May 2, 2016).

[76] ROTHSTEIN ET AL., *supra* note 35, at 24. The odds may be closer to one in trillions or even less. Ball, *supra* note 73, at 56.

[77] *Id.*

make any inferences about what the person knows or even looks like. If the fingerprint was on a murder weapon, the fingerprint points to its owner as the suspect. However, simply having someone's fingerprint without any context or the person's testimony is not helpful in determining what happened or what a person knows.

Hash values can also function like bates numbers[78] in e-discovery and may be added to documents by counsel to serve as identification.[79] It is so improbable that two files would have identical hashes,[80] that they are more reliable than fingerprints or DNA evidence.[81] Removing a single character, even something as insignificant as a period or a "space," will completely change the hash number.[82] Even opening a word document, copying all of the content, pasting it into a blank document, and saving it as a new file will produce unique files: the metadata will be different, which will generate different hash values.[83] However, changing a file name may not affect the hash for a computer file because, as discussed above, file names are not stored within the file's content but rather in the system metadata.[84] Since hash values are generated by mathematical equations, the hash value will be the same anytime a hash value is generated from a file. Thus, an identical file stored on two different computers will always have the same hash value.

Even though hash numbers are derived from the entirety of a file, the hash numbers cannot be reverse-engineered to recreate the original file.[85] If you

[78] "Bates stamping is the process of applying a set of identifying numbers to a document collection of PDFs to label and identify them." *About Bates Stamping*, LEXISNEXIS, http://help.lexisnexis.com/litigation/ac/cm/cm10/cm_bates_stamp_about.htm (last visited June 30, 2016).

[79] Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 547 (D. Md. 2007); ROTHSTEIN ET AL., *supra* note 35.

[80] Experiments with supercomputers have been able to artificially create two unique documents that have identical hash numbers. *What Are Hashes*, *supra* note 75.

[81] Ball, *supra* note 73, at 56–57.

[82] *Id.* The Hash number for Lincoln's Gettysburg address is E7753A4E97B962B36F0B2A7C0D0DB8E8; however if you change the opening line from "four score and seven years ago" to "four score and years ago" (omitting "seven") the hash number for the speech changes to 8A5EF7E9186DCD9CF618343ECF7BD00A. *Id.* If you think of hashes as dynamic locator numbers for library books, then making even the tiniest mark in the book would actually change the locator number and not just the book. Craig Ball, *A Hash of It*, BALL IN YOUR CT.: MUSINGS ON E-DISCOVERY & COMPUT. FORENSICS (Mar. 5, 2012), http://ballinyourcourt.wordpress.com/2012/03/05/a-hash-of-it/.

[83] *Id.*

[84] Ball, *supra* note 73, 56–57. Typically file names are stored in the system metadata in a separate metadata file that is not part of the file itself.

[85] *What Are Hashes*, *supra* note 75. For this reason hashes are often used as part of passwords. You create a password, which the computer translates into a hash number. When you enter the password on the computer, your entry is translated into a hash number and compared with the hash the computer created when you first created the password. If they match, you can

give someone a hash number for a computer file, you have not exchanged any content.[86] When sensitive documents are at issue and a case turns on whether someone has possession of a specific file, exchanging hashes instead of actual computer files reduces or even eliminates the need to exchange the sensitive information. For example, if a former CIA employee were accused of having classified files for which he was not authorized, lawyers can swap hashes of files on the former employee's computer and the classified files in question. That way, the CIA need not release sensitive files to the employee's lawyers to determine if the employee possesses them; secrecy is maintained if the employee is innocent or the lawyer lacks security clearance. The same is true for sensitive corporate files.

II. METADATA AND THE LAW

Metadata have been an issue in patent litigation since at least 2001[87] but in recent years has gained greater importance in litigation generally.[88] The Rules of Evidence, the Rules of Civil Procedure, and ethics laws may be implicated by how counsel handles metadata. This section explores these areas of law as well as proposed changes that may affect metadata in litigation.

A. Federal Rules of Evidence

Though ESI has certain unique properties, like all evidence, it must pass over admissibility hurdles before the trier of fact can consider it. Clearing those hurdles can pose unique challenges for electronic information for reasons that are fairly intuitive. Just think about the images we see in advertisements every day; the camera does not lie, but the software does.[89] In

proceed. If not, then you are denied access. Since the hashes cannot be reverse engineered to create the password, stealing the hashes will not grant you access to someone's computer. *See Id.*

[86] Ball, *supra* note 73, at 56–57. This can enable the identification of confidential documents without having to give the other side the documents themselves and potentially compromise their confidentiality. For example, in a case involving the theft of confidential or classified files, the party claiming the theft can provide the hashes for the files. Looking for files with the same hash on the alleged thief's computer would reveal if any sensitive files were stolen without either side unnecessarily turning over confidential files to the other. *Id.*

[87] Netword, LLC v. Centraal Corp., 242 F.3d 1347, 1353–54 (Fed. Cir. 2001) (discussing metadata and how it relates to the organization of a computer system).

[88] A search of Westlaw's case database between January 1 and December 31, 2014 reveals 188 cases addressing metadata.

[89] *See, e.g.*, Michelle Ward, *Model Meaghan Kausman Speaks Out Against Her Photoshopped Images*, PEOPLE (Aug. 28, 2014, 06:10 PM),

a digital age where anyone with a bit of computer literacy can generate or doctor files, images, and documents to look like anything, how does one prove that digital information is "what the proponent claims it is"?[90]

Before any evidence is admitted, a court must determine that the evidence is (1) relevant, (2) authentic, (3) reliable, (4) in the proper form, and (5) more probative than prejudicial.[91] Probative value is generally a case-specific issue and need not be discussed here.

1. Relevancy

Over the last decade, there has been a great deal of debate about whether metadata are or should be considered relevant to the courts.[92] One group that has been particularly influential in the debate is the Sedona Conference, a non-profit organization made up of lawyers, judges, and academics, which has issued myriad publications and recommendations for how courts, parties, and practitioners should handle electronic information.[93]

In their first publication about electronic discovery in 2004, the Sedona Conference recommended against producing or even preserving metadata in most cases. "Unless it is material to resolving the dispute, there is no obligation to preserve and produce metadata absent agreement of the parties or order of the court."[94] In short, the Conference recommended that there be a "modest legal presumption" against preserving or producing metadata.[95]

Early cases that addressed metadata's relevancy were split on whether or not to follow the Conference's recommended presumption.[96] A number of cases adopted the presumption against production[97] while others ruled that

http://www.people.com/article/model-meaghan-kausman-speaks-out-photoshopped-bikini-pic-fella-swim-australia.

[90] FED. R. EVID. 901(a) ("To satisfy the requirement of authenticating or identifying an item of evidence, the proponent must produce evidence sufficient to support a finding that the item is what the proponent claims it is.").

[91] *See, e.g.*, Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 538 (D. Md. 2007).

[92] For a more complete history, see Wescott, *supra* note 52.

[93] THE SEDONA CONFERENCE, https://thesedonaconference.org (last visited May 15, 2016).

[94] THE SEDONA CONFERENCE WORKING GRP. ON BEST PRACTICES FOR ELEC. DOCUMENT RETENTION & PROD., THE SEDONA PRINCIPLES: BEST PRACTICES RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION i (Jonathan M. Redgrave et al. eds., Jan. 2004) [hereinafter SEDONA PRINCIPLES], https://thesedonaconference.org/download-pub/99.

[95] *Id.*

[96] Wescott, *supra* note 52, at 8–14.

[97] *Id.* at 12.

metadata are presumptively relevant and should be included as part of production.[98]

In 2007, the Sedona Conference issued a revised second edition of its best practices guide. The revisions suggested something closer to presumptive relevancy of metadata. The conference recommended that documents should be produced "in the form or forms in which the information is ordinarily maintained or in a reasonably usable form."[99] The conference further explained that "taking into account the need to produce reasonably accessible metadata that will enable the receiving party to have the same ability to access, search, and display the information as the producing party."[100] The revised recommendations reflected an understanding that metadata are important to verify the authenticity of electronic documents.[101] Indeed, authenticity is, perhaps, the most important function metadata can play in litigation.[102]

2. Authenticity

Federal Rule of Evidence 901(a) states that "[t]o satisfy the requirement of authenticating or identifying an item of evidence, the proponent must produce evidence sufficient to support a finding that the item is what the proponent claims it is."[103] Metadata ultimately offer the most reliable way of determining a file's authenticity. The volume of computer data that may be produced during litigation can make traditional methods of authentication impractical, if not impossible. As one judge put it:

[98] *Id.* at 13; *see also* Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 652 (D. Kan. 2005) ("Based on these emerging standards, the Court holds that when a party is ordered to produce electronic documents as they are maintained in the ordinary course of business, the producing party should produce the electronic documents with their metadata intact, unless that party timely objects to production of metadata, the parties agree that the metadata should not be produced, or the producing party requests a protective order.").

[99] THE SEDONA CONFERENCE WORKING GRP. ON ELEC. DOCUMENT RETENTION & PROD., THE SEDONA PRINCIPLES: SECOND EDITION BEST PRACTICES RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION ii (Jonathan M. Redgrave et al. eds., June 2007) [hereinafter SEDONA PRINCIPLES SECOND EDITION] (emphasis added), https://thesedonaconference.org/download-pub/81.

[100] *Id.*

[101] *Id.* at 61 ("In assessing preservation, it should be noted that the failure to preserve and produce metadata may deprive the producing party of the opportunity later to contest the authenticity of the document if the metadata is [sic] material to that determination.").

[102] JOHN ISAZA, ARMA INT'L EDUC. FOUND., METADATA IN COURT: WHAT RIM, LEGAL AND IT NEED TO KNOW 5 (2010), http://www.armaedfoundation.org/pdfs/Isaza_Metadata_Final.pdf.

[103] FED. R. EVID. 901(a).

> "How do you demonstrate that an e-mail that you have now printed out is authentic? You may need to get the metadata to demonstrate where it came from, what its genesis was, and what its path was throughout a particular organization, in order to make your admissibility argument at trial. *So there is an argument to be made that all of that metadata is critical to the authenticity issue*."[104]

Printouts of a Word document, or a spreadsheet, could be authentic representations of documents made five years ago or five minutes ago. The printout does not necessarily tell you who created a document, where they created it, or when they created it. In cases where those facts are critical, it is worth knowing more than whose name is printed in Times New Roman at the bottom of a page. If a document has an auto-update setting for the date at the top, the printed version could be dated today even though the digital file was created years ago; moreover, opening that document updated the date, thus changing the file and changing its hash value.

Metadata authenticates a file by providing information about the context in which the file was created. A hard copy of an electronic document or an imaged version of the document strips all of the information that may answer questions about the origin of a document out of the file. Rather than relying on metadata that definitively tells you the date and time of the file's creation, the identity of the creator, and the software program that created it, you are forced to rely on other, potentially less reliable, methods of authentication.

Consider the amount of computer information a business or company generates every day. An employee's sworn statement attesting to the authenticity of a document, for example, may be countered by metadata that show the document was produced at a different time, by a different user, or in a different location than the testimony. Without sufficiently conclusive metadata, key documents may not survive an admissibility challenge under the rules of evidence.[105] So far, it does not appear that courts have addressed this issue. However, it is only a matter of time.

The nature of ESI makes it impractical, perhaps even impossible, to compile a complete list of methods to authenticate computer information.[106] However, there are a few where metadata are potentially important. In some cases, failing to provide enough evidence to authenticate ESI can lead to the exclusion of evidence. Such exclusions can determine the outcome of cases or motions. For example, in *Lorraine v. Markel American Ins. Co.*, the district

[104] Lee H. Rosenthal & James C. Francis IV, *Managing Electronic Discovery: Views from the Judges*, 76 FORDHAM L. REV. 1, 22 (2007) (emphasis added).

[105] *See* FED. R. EVID. 901(a)–(b).

[106] Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 553–54 (D. Md. 2007).

court, after a lengthy analysis of the rules of evidence, dismissed both parties' motions for summary judgment because they both failed to provide enough information to authenticate the ESI they relied on in their motions.[107] *Lorraine,* was a contractual dispute.[108] The agreement was ambiguous on its face, and each party submitted motions for summary judgment supported by printouts of e-mails and other electronic documents related to the arbitration agreement that gave rise to the dispute.[109] Neither party offered any affidavits, testimony, or metadata to show the authenticity of the documents supporting their motions.[110] As a result, the court dismissed the motions without prejudice.[111] While the *Lorraine* court considered affidavits sufficient to authenticate the files, an alternative would be to provide metadata to verify the authenticity of computer files.

The Federal Rules of Evidence provide several methods of authenticating evidence.[112] A common method of authenticating ESI is providing an affidavit or witness attesting to the authenticity from someone with knowledge of the creation of the specific document or someone with knowledge of how the type of document is routinely created.[113] The court needs "factual specificity about the process by which the electronically stored information is created, acquired, maintained, and preserved without alteration or change, or the process by which it is produced if the result of a system or process that does so."[114] Alternatively, a witness or the trier of fact can authenticate a document by comparing it to a similar but already authenticated document.[115] This technique has been used to authenticate email on at least one occasion.[116]

Authenticity can also be established based upon a document's "appearance, contents, substance, internal patterns, or other distinctive characteristics of the item, taken together with all the circumstances."[117] This method is a common means of authenticating electronic communications, particularly email.[118] The rule is notable for its reliance on circumstantial

[107] *Id.* at 585.

[108] *Id.* at 536.

[109] *Id.* at 535.

[110] *Id.* at 535–36.

[111] *Id.* at 537, 585.

[112] For simplicity, my analysis focuses on the federal rules; the analysis may vary in state courts.

[113] FED. R. EVID. 901(b)(1); *Lorraine*, 241 F.R.D. at 553–54.

[114] *Lorraine*, 241 F.R.D. at 545.

[115] FED. R. EVID. 901(b)(3).

[116] *Lorraine*, 241 F.R.D. at 546.

[117] FED. R. EVID. 901(b)(4).

[118] *Lorraine*, 241 F.R.D. at 546.

information to authenticate documents.[119] The hash numbers discussed above are particularly useful under this analysis.[120] Since each file has a unique hash based on its content, the authenticity or reliability of the file could be established based on the hash number.[121] Depending on the circumstances, information like "a file's name, a file's location . . . , file format or file type, file size, file dates . . . , and file permissions" may be enough.[122]

Lastly,[123] ESI can be authenticated by presenting "[e]vidence describing a process or system and showing that it produces an accurate result."[124] The rule was specifically designed to accommodate computer files and systems where the system assumes or depends on the accuracy of the information presented.[125] For example, if the question is whether a particular employee was at her desk at a particular time, one party can produce evidence that the employee was logged into her computer on the date and time at issue or that certain network features were accessed from that computer within a couple of minutes of the time at issue. The proponent of the evidence then can demonstrate reliability by showing that their network monitoring is important for cyber security or by showing that if it were inaccurate, the system would not function. Thus, the metadata about who is using what computer system on the network and at what time are reliable because the business depends on the accuracy of the information.

This last method of authentication is closely related to another evidentiary hurtle for ESI, reliability.

3. Reliability

The most common question in reliability is hearsay. However, hearsay analyses are largely inapplicable to metadata and other information generated by a computer based on users' input, "[b]ecause such records are not the counterpart of a statement by a human declarant[,] . . . they should not be treated as hearsay[.] [B]ut rather their admissibility should be determined on

[119] *Id.*

[120] *See supra* part I.B.4.

[121] *Lorraine*, 241 F.R.D. at 546–47; *see also* PAUL W. GRIMM & LISA M. YURWIT, ELECTRONICALLY STORED INFORMATION IN MARYLAND AND FEDERAL COURTS: DISCOVERY, ADMISSIBILITY, AND ETHICS CHAPTER NINE: ADMISSIBILITY—RELEVANCE, AUTHENTICITY, HEARSAY, AND THE ORIGINAL WRITING RULE (2008).

[122] *See* Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 646 (D. Kan. 2005).

[123] Authentication of public records can also depend on electronic information. For a more detailed discussion, see *Lorraine*, 241 F.R.D. at 547–49.

[124] FED. R. EVID. 901(b)(9).

[125] *See* FED. R. EVID. 901 ex. 9; *see also Lorraine*, 241 F.R.D. at 548.

the basis of the reliability and accuracy of the process involved."[126] In essence, "nothing 'said' by a machine . . . is hearsay."[127]

However, admissibility of a statement under certain hearsay exceptions can turn on metadata.

4. Exceptions to the Rule Against Hearsay

The Federal Rules of Evidence recognize twenty-nine exceptions to the rule against hearsay.[128] There is also a catchall exception for other statements that "ha[ve] equivalent circumstantial guarantees of trustworthiness."[129] Determining whether the statements meet a hearsay exception may be best determined by an analysis of metadata. Electronic information is implicated in many of these hearsay exceptions, and the growing popularity of social media, instant messaging, text messages, and other electronic means of communication increases the reliance on electronic information in these analyses. For example, present sense impressions or excited utterances[130] may

[126] United States v. Rollins, No. ACM34515, 2004 WL 26780, at *10 (A.F. Ct. Crim. App. Dec. 24, 2003) (quoting State v. Dunn, 7 S.W.3d 427, 432 (Mo. Ct. App. 1999) (internal quotations omitted)), *aff'd in part, rev'd in part on other grounds and remanded*, 61 M.J. 338 (C.A.A.F. 2005).

[127] United States v. Khorozian, 333 F.3d 498, 506 (3d Cir. 2003) (quoting 4 CHRISTOPHER MUELLER & LAIRD KIRKPATRICK, FEDERAL EVIDENCE § 380, at 65 (2d ed.1994)); *see also Dunn*, 7 S.W.3d at 431 ("A trace report, which tracks a telephone call made to a specific number and which is generated by a telephone company's computer, is not hearsay . . . . Because records of this type are not the counterpart of a statement by a human declarant, which should ideally be tested by cross-examination of that declarant, they should not be treated as hearsay, but rather their admissibility should be determined on the basis of the reliability and accuracy of the process involved.").

[128] Regardless of whether the witness is available: present sense impressions; excited utterances; then existing mental, emotional, or physical condition; statements made for medical diagnosis and treatment; recorded recollections; records of regularly conducted activities; absence of a record of a regularly conducted evidence; public records; public records of vital statistics; absence of a public record; records of religious organizations concerning personal or family history; certificates of marriage, baptism, and similar ceremonies; family records; records of documents that affect an interest in property; statements in documents that affect an interest in property; statements of ancient documents; market reports and similar commercial publications; statements in learned treatises, periodicals, or pamphlets; reputation concerning personal or family history; reputation concerning boundaries or general history; reputation concerning character; judgment of previous conviction; judgments involving personal, family, or general history, or a boundary. FED. R. EVID. 803. If the declarant is unavailable: statements made under the belief of imminent death; statements against interest; statements of personal or family history; statement against the party that wrongfully cause the declarant's unavailability. FED. R. EVID. 804(b).

[129] FED. R. EVID. 807(a)(1).

[130] FED. R. EVID. 803(2).

take the form of a text message,[131] tweet, Facebook post, or email rather than a spoken utterance to another person.[132]

Since metadata are created as part or byproduct of human interactions with electronic devices, they can provide evidence to show whether a statement is subject to a hearsay exception.

a. Business Records Exception

The Federal Rules of Evidence allow records kept in the ordinary course of business that reflect a regular activity of the business in question to be admitted even if the content would normally be excluded as hearsay.[133] Electronic information raises questions of what is considered "regularly conducted activity." E-mails in particular can be both routine and extraordinary depending on the content and context.

Courts have taken a range of approaches to permitting ESI under this exception.[134] At one extreme, courts require a detailed showing of the companies practice for email generation and retention.[135] For example, in *New York v. Microsoft*, the D.C. District Court conducted a lengthy analysis of e-mails sent between employees to determine if they were records prepared in the normal course of business.[136] The court ultimately concluded that even though the employee routinely sent e-mails to customers following a phone call, an employee's habit was not enough to prove that those e-mails were business records for the purposes of the hearsay exception.[137] The court was particularly concerned that the evidence did not show that it was the company's policy for the e-mails to be transmitted rather than the habit of a

[131] State v. Damper, 225 P.3d 1148, 1152 (Ariz. Ct. App. 2010) (upholding a trial court's ruling that text messages from a murder victim to a third party were present sense impressions, because they included present tense statements like "Me and Marcus are fighting").

[132] Some commentators have argued that these exceptions should not apply to electronic media. Jeffrey Bellin, *Facebook, Twitter, and the Uncertain Future of Present Sense Impressions*, 160 U. PA. L. REV. 331, 362 (2012).

[133] FED. R. EVID. 803(6)(A)–(C), (E) ("[T]he record was made at or near the time by—or from information transmitted by—someone with knowledge; [] the record was kept in the course of a regularly conducted activity of a business, organization, occupation, or calling, whether or not for profit; [] making the record was a regular practice of that activity . . . [neither] the source of information [n]or the method or circumstances of preparation indicate a lack of trustworthiness.").

[134] Grimm & Yurwit, *supra* note 121.

[135] Rambus, Inc. v. Infineon Techn. AG, 348 F. Supp. 2d 698, 707 (E.D. Va. 2004) (ruling on the admissibility of e-mail chains as business records when one of the mails came from someone outside the company); New York v. Microsoft Corp., No. 98-1233 (CKK), 2002 U.S. Dist. LEXIS 7683, at *8–9 (D.D.C. Apr. 12, 2002).

[136] *Microsoft Corp.*, 2002 U.S. Dist. LEXIS 7683, at *9.

[137] *Id.*

particular employee.[138] The court noted the complete lack of evidence presented by the parties about the company's normal practice for generating and retaining e-mails; absent such a showing, the court found that "the method or circumstances of preparation indicate lack of trustworthiness."[139]

On the other end of the spectrum, the same district court in *U.S. v. Safavian* determined that similar objections to the authenticity of e-mail chains were questions of weight more properly addressed by the jury and did not pertain to the authenticity or reliability of the e-mails under a hearsay analysis.[140]

The distinct approaches taken by these two district courts show that the admissibility of electronic evidence under the business records exception may depend on evidence of how the records were created. Such determinations may come down to the information contained in the metadata of communications. In *Safavian*, the court looked to the e-mail addresses in the e-mails to authenticate them.[141] The court's analysis focused on whether the e-mail addresses themselves included the senders' names or something sufficiently close to them to determine if the e-mails were transmitted between the individuals alleged.[142]

Metadata can be helpful in analyses like this. If a document was generated in the ordinary course of business, then the metadata about the file's creation date, creator, format, etc. should match other similar documents. Discrepancies in that data can demonstrate that it was not actually generated in the course of business and should be subject to a traditional hearsay analysis. As with its other uses in litigation, metadata can provide useful information about the circumstances surrounding a document's creation and history within a computer server. Indeed, in *Stallings v. City of Johnson*, the District Court accepted this argument and permitted expert testimony from

[138] *Id.*

[139] *Id.* at *9 (quoting FED. R. EVID. 803(6)) ("While Mr. Glaser's email may have been 'kept in the course' of RealNetworks regularly conducted business activity, Plaintiffs have not, on the present record, established that it was the 'regular practice' of RealNetworks employees to write and maintain such emails."); *see also id.* at *14 ("If both the source and the recorder of the information, as well as every other participant in the chain producing the record, are acting in the regular course of business, the multiple hearsay is excused by Rule 803(6).") (quoting United States v. Baker, 693 F.2d 183, 187 (D.C. Cir. 1982)).

[140] United States v. Safavian, 435 F. Supp. 2d 36, 41 (D.D.C. 2006) ("The defendant argues that the trustworthiness of these e-mails cannot be demonstrated, particularly those e-mails that are embedded within e-mails as having been forwarded to or by others or as the previous e-mail to which a reply was sent. The Court rejects this as an argument against authentication of the e-mails. The defendant's argument is more appropriately directed to the weight the jury should give the evidence, not to its authenticity.").

[141] *Id.* at 40–41.

[142] *Id.*

the creator of a software program to explain how specific metadata were produced and how data was kept in the ordinary course of business.[143]

However, the ease with which ESI, metadata in particular, can be modified raises questions about how to produce the original version of a document.

b. The Original Document Rule

While the Original Document Rule makes intuitive sense in the physical world, it is difficult to apply to electronic information. The Federal Rules of Evidence lay out a series of circumstances where the original document is required and a duplicate will not suffice. As a general matter, the original document is required when the contents of the document are in question.[144] Unless there is a question about the authenticity of an original document, a duplicate can be admitted instead.[145] Notably, "[f]or electronically stored information, 'original' means any *printout*—or other output readable by sight—if it accurately reflects the information."[146]

The rules state and are interpreted to mean that a computer printout or a copy of an electronic document is sufficient;[147] in other words, no metadata are needed. However, courts are starting to question this principle. The Third Circuit Court of Appeals recently noted that determining what constitutes an original document is growing increasingly difficult.

> Moving from the more easily distinguishable photocopy or facsimile to documents created, transmitted and stored in an electronic form means that it will be increasingly difficult to ascertain where the boundary of an objectively reasonable duty to preserve such documents lies. *There are—and increasingly will be—circumstances in which the foreseeability of a duty to preserve the information contained in a particular document is distinguishable—under an objective analysis—from the need to*

[143] *See cf.* Stallings v. City of Johnson City, No. 13-cv-422-DRH-SCW, 2016 WL 424819, at *2–3 (S.D. Ill. Feb. 4, 2016).

[144] FED. R. EVID. 1002 ("An original writing, recording, or photograph is required in order to prove its content."); *see also* Rebecca Levy-Sachs & Taylor Archambault, *Hurdling Toward the Future: Navigating Electronically Stored Information Through the Federal Rules of Evidence:* Lorraine v. Markel America Insurance Co., *in* EVIDENCE ESI 6, 11–12 (Ralph A. Zappala & Rebecca Levy-Sachs eds., 2008), http://www.thefederation.org/documents/10.LevySachs.pdf ("To summarize the Court's guidance, when offering ESI as evidence, practitioners should make the threshold determination of whether the original writing rule applies and then be prepared to produce an original, a duplicate, or secondary evidence of the contents."); *see generally* Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 576–83 (D. Md. 2007).

[145] FED. R. EVID. 1003.

[146] FED. R. EVID. 1001(d) (emphasis added).

[147] *Lorraine*, 241 F.R.D at 577–78.

> *preserve that information in its "original" form or format*. Indeed, arriving at a common understanding of what an "original" is in this context is challenging enough. Although it does, and always will rest with the courts to preserve the distinction between an objectively foreseeable duty and actual knowledge of such a duty, there is a concomitant obligation that counsel must assume to clearly and precisely articulate the need for parties to search for, maintain, and—where necessary—produce "original" or source documents.[148]

While the court did not explicitly mention metadata, the discussion clearly implicates the dangers of intentionally, or even inadvertently, destroying metadata. If identifying information is removed, a document may no longer be in its original form. Indeed, a duplicate of an electronic document can easily appear on screen or in printed form exactly like the original. However, the duplicate will have different metadata identifying a different creation or modification date, owner, file location, and more.[149]

The basic principle of the rule is that any medium that accurately displays the *content* of the original electronic document is considered an original for purposes of the Federal Rules of Evidence.[150] However, lawyers may need to present e-mails in their original electronic form, and not a computer printout if their authenticity is at issue.[151] Generally speaking, lawyers should determine if the Original Document Rule applies and then produce additional evidence to demonstrate the reliability and authenticity of the document.[152] In these analyses, the metadata can show that the documents presented to the court are the emails that are at issue in the dispute.[153]

Though the current rules of evidence do not on their face consider metadata a part of the original document, courts have begun to make exceptions for documents like emails.[154] The circumstances of a specific case

[148] Bull v. United Parcel Serv., Inc., 665 F.3d 68, 78 n.12 (3d Cir. 2012) (emphasis added).

[149] *See supra* note 59 and accompanying text.

[150] *See* Laughner v. State, 769 N.E.2d 1147, 1159 (Ind. Ct. App. 2002) (ruling that computer printouts of internet chat rooms were admissible as evidence to prove child sex solicitation) *abrogated by* Fajardo v. State, 859 N.E.2d 1201 (Ind. 2007) (not addressing the original evidence rule).

[151] *Lorraine*, 241 F.R.D. at 583 ("In this case, counsel did not address the original writing rule, despite its obvious applicability given that the e-mail exhibits were closely related to a controlling issue and there [sic] were proving the contents of the e-mails themselves.").

[152] *See also* Levy-Sachs & Archambault, *supra* note 144; *see generally Lorraine*, 241 F.R.D. at 576–83.

[153] *See Lorraine*, 241 F.R.D. at 583, 585 (dismissing motions for summary judgment because the parties failed to apply the original document rule to e-mails).

[154] *See* Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 651 (D. Kan. 2005); Aguilar v. Immigration & Customs Enf't Div. of U.S. Dep't of Homeland Sec., 255 F.R.D. 350, 359

may require production of information other than the pure content of a file. While these questions await a more definitive resolution, practitioners run the risk of destroying evidence by not preserving metadata.

B. Discovery and Spoliation

When the Sedona Conference made its initial recommendations in 2004 that there be a "modest legal presumption" against the relevancy and discoverability of metadata,[155] it was advocating a position contrary to the plain meaning of the contemporary Federal Rules of Civil Procedure. At that time, Rule 34 stated, "a party who produces documents for inspection shall produce them as they are kept in the *usual course of business*."[156] Under the plain meaning of the rule, producing a document without the metadata violated the rule.[157] However, some courts felt the guidance in the rules, including the then proposed amendments, failed to answer the question.[158]

1. Early Decisions on ESI and Metadata

Around the same time as the Sedona Conference recommendations, and equally, if not more, influential in cases addressing ESI, were the *Zubulake* decisions in 2004.[159] *Zubulake v. UBS Warburg LLC* started as a "garden variety employment discrimination case."[160] After Zubulake filed an EEOC complaint alleging gender discrimination in promotion decisions and

(S.D.N.Y. 2008) (refusing to require the production of documents with the "BCC" information intact because the request was made after production was mostly complete).

[155] SEDONA PRINCIPLES, *supra* note 94, at 41.

[156] FED. R. CIV. P. 34(b) (2004) (emphasis added).

[157] *See* Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 654 (D. Kan. 2005) ("[T]aken in the context of Plaintiffs' stated reasons for requesting the Excel spreadsheets in their native electronic format and the Court's repeated statements that the spreadsheets should be produced in the electronic form in which they are maintained, the Court finds that Defendant should have reasonably understood that the Court expected and intended for Defendant to produce the spreadsheets' metadata along with the Excel spreadsheets.").

[158] *Id.* at 649 ("Although the proposed amendments to Rule 34 use the phrase 'in a form or forms in which it is ordinarily maintained,' they provide no further guidance as to whether a party's production of electronically stored information 'in the form or forms in which it is ordinarily maintained' would encompass the electronic document's metadata.").

[159] Zubulake v. UBS Warburg LLC (*Zubulake V*) , 229 F.R.D. 422 (S.D.N.Y. 2004). These were a series of five decisions issued in the course of the discovery phase of a single case before Judge Scheindlin in the Southern District of New York. The parties fervently disagreed about the defendant's duty to preserve and produce ESI.

[160] Victor Li, *Looking Back on Zubulake, 10 Years Later*, ABA JOURNAL (Sept. 1, 2014, 10:00 AM), http://www.abajournal.com/magazine/article/looking_back_on_zubulake_10_years_later.

harassment in the workplace, her employer fired her.[161] Zubulake then sued her employer in federal court alleging retaliation for reporting to the EEOC.[162]

Zubulake's attorneys realized they were not receiving all of the documents they needed in response to discovery requests.[163] The principle issue was UBS's back-up tapes for their e-mail system.[164] The tapes could not be searched; the only way to view the information was to restore the information onto a server, a lengthy and expensive procedure.[165] In the end, one of the relevant back-up tapes was overwritten before a final determination was made about restoring it; Judge Scheindlin gave an adverse inference instruction to the jury as a sanction for failing to preserve evidence.[166] This meant that even though the jury could not actually view the original evidence, they were instructed to presume that the contents of the tape supported Zubulake's claims.[167] Though it may not have affected *Zubulake*'s outcome,[168] such an adverse inference could have a major impact on litigation in other contexts. The effect of this decision was immediate and broad; lawyers and corporations were on notice that failing to preserve ESI could result in harsh sanctions in litigation.

The *Zubulake* decisions and the 2006 FRCP Amendments have effectively defined attorneys' and parties', particularly corporate parties',[169] obligation to preserve information during and in preparation for litigation. As Judge Scheindlin stated, "[t]he obligation to preserve evidence arises when the party has notice that the evidence is relevant to litigation or when a party should have known that the evidence may be relevant to future litigation."[170] The

[161] *Id.*

[162] *Id.*

[163] *Id.* ("The tip-off for Zubulake's employees that they weren't getting all of the electronic evidence available from UBS came when . . . [the firm] produced only 120 emails in response to Zubulake's document request. . . . on her own, she [Zubulake] had printed out more than 400 emails that were relevant to the complaint.").

[164] Zubulake v. UBS Warburg LLC (*Zubulake I*), 217 F.R.D. 309, 313 (S.D.N.Y. 2003).

[165] *Id.*

[166] Zubulake V, 229 F.R.D. 439–40 (S.D.N.Y. 2004).

[167] *Id.* at 437.

[168] Li, *supra* note 160.

[169] *See generally* THE SEDONA CONFERENCE WORKING GRP. ON BEST PRACTICES FOR ELEC. DOCUMENT RETENTION & PROD., THE SEDONA GUIDELINES: BEST PRACTICE GUIDELINES & COMMENTARY FOR MANAGING INFORMATION & RECORDS IN THE ELECTRONIC AGE (Lori Ann Wagner ed., 2nd ed. Nov. 2007) [hereinafter THE SEDONA GUIDELINES], https://thesedonaconference.org/publication/Managing%20Information%20%2526%20Records.

[170] Zubulake v. UBS Warburg LLC (*Zubulake IV*), 220 F.R.D. 212, 216 (S.D.N.Y. 2003) (internal citations and quotations omitted).

duty to preserve evidence lands first on the lawyer who must inform the client of the obligation to preserve potentially relevant evidence.[171]

In *Zubulake V* the court noted that "[c]ounsel must take affirmative steps to monitor compliance so that all sources of discoverable information are identified and searched . . . . [C]ounsel and client must take *some reasonable steps* to see that sources of relevant information are located."[172] In some cases, counsel may need to take possession of digital evidence to ensure it is not destroyed.[173] The *Zubulake* case led to a massive expansion of the e-discovery industry and spawned companies who specialized in the retrieval, analysis, and preservation of electronic information.[174] In this regard, it is hard to understate the effect these opinions had on litigation and discovery.[175]

Based on these opinions and the subsequent changes to the Rules of Civil Procedure, lawyers' duties in discovery are now generally defined by the standards Judge Scheindlin created. First, lawyers should determine what their client's document retention policies are—or create one if no such policy exists.[176] Once litigation begins or is reasonably foreseeable, attorneys should issue litigation hold letters instructing their clients not to delete any relevant documents.[177] As part of the process, lawyers must learn their client's computer systems and back-up policies and actively ensure that everyone involved knows what must be preserved.[178]

It is difficult to say when a lawyer is required to take possession of a client's computer files to ensure their preservation. However, the facts of *Zubulake* indicate that when there is a substantial likelihood that a client's normal practice, such as overwriting back-up tapes at regular intervals, makes the destruction of evidence likely, a lawyer should take possession of the electronic information to ensure its preservation. Though the *Zubulake*

[171] Telecom Int'l Am., Ltd. v. AT & T Corp., 189 F.R.D. 76, 81 (S.D.N.Y. 1999).

[172] *Zubulake V*, 229 F.R.D. at 432.

[173] *See id.* at 434.

[174] Li, *supra* note 160.

[175] Perhaps realizing the effect the opinions would have, Judge Scheindlin concluded the *Zubulake V* opinion by saying, "[n]ow that the key issues have been addressed and national standards are developing, parties and their counsel are fully on notice of their responsibility to preserve and produce electronically stored information. The tedious and difficult fact finding encompassed in this opinion and others like it is a great burden on a court's limited resources. The time and effort spent by counsel to litigate these issues has also been time-consuming and distracting. This Court, for one, is optimistic that with the guidance now provided it will not be necessary to spend this amount of time again. It is hoped that counsel will heed the guidance provided by these resources and will work to ensure that preservation, production and spoliation issues are limited, if not eliminated." *Zubulake V*, 229 F.R.D. at 440–41.

[176] Perkins & Deppe, *supra* note 72.

[177] *Id.*

[178] *Id.*

decisions did not address metadata, Judge Scheindlin's reasoning could apply to metadata, particularly when the facts of a case or the nature of the evidence hinges on authenticating electronic information or identifying online activities of a party.

2. The 2006 Amendments to the Federal Rules of Civil Procedure

Following the *Zubulake* decisions and in response to the confusion over the original document rule, the Federal Rules of Civil Procedure were amended in 2006. The revised rules create a framework for courts and lawyers to handle electronic discovery. The rules broadly define the types of ESI that may be discoverable or requested by a party.[179] The rules also permit sampling or testing so that vast numbers of documents are not produced unnecessarily.[180] The revised rules retain the "usual course of business" language,[181] which leaves open the issue of metadata.[182]

As a result, there is no clear rule for the metadata's production. Documents kept in the usual course of business will almost certainly include metadata, and courts have generally interpreted the "usual course of business" to mean documents' "native format,"[183] including the metadata.[184]

[179] FED. R. CIV. P. 34(a)(1)(A) ("A party may serve on any other party a request . . . to produce and permit the requesting party or its representative to inspect, copy, test, or sample the following items in the responding party's possession, custody, or control [] any designated documents or electronically stored information—including writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations—stored in any medium from which information can be obtained either directly or, if necessary, after translation by the responding party into a reasonably usable form[.]").

[180] FED. R. CIV. P. 34(a)(1).

[181] FED. R. CIV. P. 34(b)(2)(E)(i) ("[a] party must produce documents as they are kept in the usual course of business or must organize and label them to correspond to the categories in the request").

[182] The amended rules' only mention of metadata relates to waiver and leaves the issues to the discretion of the parties and the court. *See* FED. R. CIV. P. 26(f) (advisory committee's note to 2006 amendment) ("Information describing the history, tracking, or management of an electronic file (sometimes called "metadata") is usually not apparent to the reader viewing a hard copy or a screen image. Whether this information should be produced may be among the topics discussed in the Rule 26(f) conference.").

[183] "'Native File(s)' means ESI in the electronic format of the application in which such ESI is normally created, viewed and/or modified. Native Files are a subset of ESI." PAUL W. GRIMM ET AL., SUGGESTED PROTOCOL FOR THE DISCOVERY OF ELECTRONICALLY STORIED INFORMATION 3 (2016) http://www.mdd.uscourts.gov/news/news/ESIProtocol.pdf.

[184] *See* THOMAS ALLMAN, THE SEDONA CONFERENCE, STATE E-DISCOVERY TODAY: AN UPDATE ON RULEMAKING IN LIGHT OF THE 2006 FEDERAL AMENDMENTS 17–18 (citations omitted), https://thesedonaconference.org/system/files/Chapter%202%20%20State%20eDiscovery%20Today.pdf.

However, there is considerable debate about metadata's relevance to discovery; federal and state courts have come to very different conclusions.[185] The general rule is that metadata are discoverable if they are relevant,[186] and it is up to the parties to decide what metadata should be produced in a dispute.[187] In sum, the current rules of civil procedure barely address metadata and leave it to the courts and parties to deal with the issue on an *ad hoc* basis.

However, parties and attorneys can still face sanctions under Rule 37 for failing to properly preserve metadata. "Absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information lost as a result of the routine, good-faith operation of an electronic information system."[188] The language led to a circuit split on when adverse inferences are appropriate; some courts require mere negligence to trigger an adverse inference while others require bad faith.[189] In addition, the Third Circuit has indicated that intentionally refusing to produce original documents could amount to spoliation.[190]

3. The 2014 Amendments to the Federal Rules of Civil Procedure

The proposed amendments to the Federal Rules of Civil Procedure do not take a stand on metadata; rather, they focus primarily on judicial management of cases, spoliation of ESI, and proportionality of discovery.[191] The spoliation provisions will cure the circuit split on adverse inferences by requiring bad

[185] *See id.*

[186] *Id.* at 20.

[187] Thomas Y. Allman, *The Impact of the Proposed Federal E-Discovery Rules*, 12 RICH. J.L. & TECH. 13, 21 (2006) [hereinafter *Impact of Proposed Rules*] ("Neither default form [of production] is intended to mandate production of metadata or embedded data."); s*ee also* Allman, *supra* note 184, at 19.

[188] FED. R. CIV. P. 37(e) (advisory committee's note to 2015 amendment). This language could even amount to broad immunity for destroying metadata, because metadata are so easily modified. Of course if a party knows of the risk, it is harder to make a good faith argument if the party does not take steps to preserve relevant metadata.

[189] Memorandum from Judge David G. Campbell on Proposed Amendments to the Federal Rules of Civil Procedure to Judge Jeffrey Sutton 17 (June 14, 2014), www.uscourts.gov/file/18218/download.

[190] Bull v. United Parcel Serv., Inc., 665 F.3d 68, 78–79 n.12 (3d Cir. 2012) (applying the spoliation factors from Brewer v. Quaker State Oil Ref. Corp., 72 F.3d 326, 334 (3d Cir. 1995) and discussing the increasing difficulty with determining when there is a duty to preserve electronic documents in their "original" form).

[191] *See generally* Memorandum from Judge David G. Campbell, *supra* note 189 ("[O]nly upon finding that the party acted with the intent to deprive another party of the information's use in the litigation, may [the court]: (A) presume that the lost information was unfavorable to the party; (B) instruct the jury that it may or must presume the information was unfavorable to the party; or (C) dismiss the action or enter a default judgment.").

faith and an intent to deprive the other party of information in discovery.[192] However, negligent destruction of evidence could still lead to other sanctions at the discretion of the court.[193]

Absent from the proposed amendments is any discussion of metadata. Presumably, metadata spoliation would fall under the new Rule 37(e)'s bad faith and negligence analyses. However, there is no discussion about presumptive relevancy or use for authentication.

C. Confidentiality and Waiver Issues

Any discussion of the obligation to preserve metadata requires understanding the types of information lawyers routinely have to produce and screen for privilege. E-discovery is an expensive[194] and time-consuming affair. A search of a party's computer systems for responsive files could bring up thousands or hundreds of thousands of documents. The process may require lawyers to take possession of a client's files or computers. Invariably lawyers will take custody of a client's files in preparing to respond to discovery requests.[195]

Taking possession of digital evidence creates new questions. For example, a number of states' ethics commissions have issued opinions about the extent to which lawyers can make files they took possession of during litigation available to their clients.[196] However, the primary concern for attorneys has

[192] *See id.* at 57 (citing to amendments of FED. R. CIV. P. 37(e)(2)).

[193] *See id.* (citing to amendments of FED. R. CIV. P. 37(e)(1)).

[194] It costs at least $3,500 for counsel to review a gigabyte of data. Dean Gonsowski, *E-Discovery Costs: Pay Now or Pay Later*, INSIDE COUNS. (May 23, 2012), http://www.insidecounsel.com/2012/05/23/e-discovery-costs-pay-now-or-pay-later. Seventy-three percent of the cost of discovery is the cost of having a lawyer review it for privilege and relevancy. NICHOLAS M. PACE & LAURA ZAKARAS, WHERE THE MONEY GOES: UNDERSTANDING LITIGANT EXPENDITURES FOR PRODUCING ELECTRONIC DISCOVERY xiv–xv (2012), http://www.rand.org/pubs/monographs/MG1208.html.

[195] California law states that client files in the lawyer's custody remain the property of the client and not the lawyer. State Bar of Cal. Standing Comm. on Prof'l Responsibility & Conduct, Formal Op. 1994-134, http://ethics.calbar.ca.gov/LinkClick.aspx?fileticket=1zWyNtvVULE%3D&tabid=839#FNT4 (last visited May 16, 2016). Under California law, a lawyer must return the client's files if requested. *Id.* Under those circumstances, returning electronic information with modified metadata could lead to questions of the authenticity of computer files; those questions could implicate the lawyer. *Id.* This is particularly true if the files are being returned so the client can retain different counsel. *Id.*

[196] State Bar of Ariz. Ethics Opinions, *09-04: Confidentiality; Maintaining Client Files; Electronic Storage; Internet* (Dec. 2009), http://www.azbar.org/Ethics/EthicsOpinions/ViewEthicsOpinion?id=704; Ala. State Bar, *Retention, Storage, Ownership, Production and Destruction of Client Files* (2010),

been, and should be, ensuring client confidentiality. "A lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent, the disclosure is impliedly authorized in order to carry out the representation . . . ."[197] However, there are exceptions available at the lawyer's discretion to prevent death or imminent harm to others, fraud, or for self-defense.[198]

It is common for lawyers, clients, and experts to exchange documents and comment or modify them, perhaps using tracked changes. These changes and exchanges are covered by both the attorney-client privilege and the work product doctrine, and thus need to be removed before a document is produced. If documents are not properly redacted, then sensitive, confidential, or even classified[199] information may be revealed inadvertently. Initial confusion over what to do about inadvertent disclosure was largely solved by the adoption of Federal Rule of Evidence 502(b), which retains confidentiality and privilege so long as the affected party made good faith efforts to maintain the privilege.[200] Similarly, the Model Rules include a requirement for the party receiving documents to notify the producing party if documents were transmitted that inadvertently pierce the attorney-client privilege.[201] Nevertheless, parties often "mine" for information in produced documents looking for additional information in the hidden or redacted portions of documents. Thus, scrubbing metadata by creating a new file or

https://www.alabar.org/resources/office-of-general-counsel/formal-opinions/2010-02/; State Bar of Nev. Standing Committee on Ethics and Prof'l Responsibility, *Formal Opinion No. 33* (Feb. 9, 2006), http://ftp.documation.com/references/ABA10a/PDfs/3_12.pdf.

[197] MODEL RULES OF PROF'L CONDUCT r. 1.6(a) (AM. BAR ASS'N 2013).

[198] MODEL RULES OF PROF'L CONDUCT r. 1.6(b) (AM. BAR ASS'N 2013).

[199] *See* Gene Koprowski, *NSA and the Dangers of Documents*, ECONTENT (Apr. 13, 2006), http://www.econtentmag.com/Articles/News/News-Feature/NSA-and-the-Dangers-of-Documents-15304.htm (copying and pasting a redacted Pentagon document removed the redactions exposing confidential information).

[200] FED. R. EVID. 502(b) ("When made in a federal proceeding or to a federal office or agency, the disclosure does not operate as a waiver in a federal or state proceeding if: (1) the disclosure is inadvertent; (2) the holder of the privilege or protection took reasonable steps to prevent disclosure; and (3) the holder promptly took reasonable steps to rectify the error, including (if applicable) following Federal Rule of Civil Procedure 26 (b)(5)(B).").

[201] MODEL RULES OF PROF'L CONDUCT r. 4.4 (AM. BAR ASS'N 2013).

imaging a document is a common and even suggested practice.[202] However, scrubbing information also means removing relevant metadata.[203]

When addressing confidentiality considerations, the distinction between data and metadata is particularly important. Scrubbing files of privileged information is when lawyers are most likely to inadvertently destroy or modify metadata.[204] As discussed above, the term "metadata" is often broadly, albeit inaccurately, defined to include any hidden information.[205] Even experts in e-discovery will use the term "metadata" to describe hidden parts of document's content.[206] The broad definition would encompass things like tracked changes, blacked out text, hidden columns, and past file versions. However, as addressed above, these are examples of data, not metadata.[207]

Nevertheless, when lawyers are scrubbing their documents of privileged data to preserve confidentiality, they are usually concerned with pseudo-metadata.[208] This data at issue in confidentiality and privilege analyses is distinctly different from true metadata that can authenticate documents. Some of the true metadata may still be privileged if they includes authors' names or past viewers that may pierce the attorney-client privilege, but this is rarely the case. It would be a mistake for a practitioner to focus only on preserving confidentiality and not also preserving metadata.

Though metadata may not have been relevant at the start of discovery, it may suddenly become relevant as more information comes to light and theories of the case evolve to match the facts. If a practitioner is focused only on protecting confidentiality, she may inadvertently destroy metadata that are

[202] *See* ABA Comm'n on Ethics & Prof'l Responsibility, Formal Op. 06-442 (2006) ("A lawyer who is concerned about the possibility of sending, producing, or providing to opposing counsel a document that contains or might contain metadata, or who wishes to take some action to reduce or remove the potentially harmful consequences of its dissemination, may be able to limit the likelihood of its transmission by "scrubbing" metadata from documents or by sending a different version of the document without the embedded information."); *see also* Burke T. Ward, et al., *Electronic Discovery: Rules for a Digital Age*, 18 B.U. J. SCI. & TECH. L. 150, 172 (2012).

[203] *See supra* section I.B.3.

[204] In cases where the metadata is necessary to establish key factual issues, scrubbing the metadata can be grounds for sanctions. *See generally* CQuest Am., Inc. v. Yahasoft, Inc., No. 13-cv-3349, 2015 WL 4576778 (C.D. Ill. July 30, 2015) (considering a motion for sanctions when the producing party produced a PDF of the change log rather than the metadata "which could have told plaintiff who made which changes and when" rather than showing all changes as the date the PDF was created).

[205] See *supra* Part II.B.

[206] *See* Wescott, *supra* note 52, at 4 (emphasis added) ("Substantive metadata, according to the *Suggested Protocol*, is 'data that reflects the *substantive changes made to the document by the user*.'").

[207] *See supra* Part II.B.

[208] I use this term to describe information that is hidden in the file but is not technically metadata. *See supra* Part II.B.

relevant. Indeed, once the documents are produced, the opposing party may want the metadata for the documents. Even if the metadata were not maliciously destroyed, missing or damaged metadata can raise eyebrows for other parties in the case and for the judge who may be mediating discovery disputes. Thus, it is important for practitioners to develop sound policies to preserve metadata when handling clients' documents.

III. LAWYERS SHOULD ALWAYS PRESERVE METADATA

Though confidentiality is a major concern for lawyers taking possession of clients' files, lawyers also need to be cognizant of the danger of damaging or altering metadata.[209] As discussed above, even a minor change to a document may alter it sufficiently to prevent its correct identification.[210] The foregoing discussion demonstrates that metadata are relevant in two key respects. First, metadata can authenticate other evidence or raise doubt about the authenticity of proffered electronic information. Metadata can also prove that a document or electronic writing of a witness meets one of the enumerated hearsay exceptions. Second, metadata can be evidence to establish, *inter alia*, notice, the location of a key witness, or the adherence to or violation of a normal procedure.

The Federal Rules of Civil Procedure allow for sanctions against parties for failing to preserve ESI. Though the newest revisions require bad faith to issue adverse inferences against a party, gentler sanctions including attorneys' fees could be awarded to a party for negligently destroying metadata. Given how susceptible metadata are to modification, the chances that a practitioner may negligently modify information are higher than the chances of negligently destroying the content of a file.

Given this uncertainty and the important role metadata can play in litigation, lawyers' best practice when dealing with electronic information is to preserve the metadata for any client files that are in their possession. Thus, if metadata become relevant evidence during litigation, they are readily available for extraction and production without requiring the client to spend time relocating documents that have already been produced. Furthermore, if there is a question about spoliation during the case, the lawyer will have the original electronic information preserved and ready for analysis. If there was

[209] This analysis is only concerned with the metadata of clients' files. However, lawyers should also preserve metadata for files that they create. *See* Long Bay Mgmt. Co. v. Haese LLC, 40 N.E.3d 1056 (Mass. App. Ct. 2015) (upholding a saction of a default judgement against a law firm sued by its former client for, *inter alia*, not producing the metadata of the client's billing file).

[210] *See supra* note 82 and accompanying text.

negligence or even intentional destruction of evidence, the lawyer will likely be quickly exculpated from any accusation of wrongdoing, saving the court and all parties a great deal of time and energy.

Thus, when lawyers receive documents from their clients, they should preserve the original document. When a lawyer receives a hard drive, flash drive, computer, or any other storage device containing client files, the lawyer should forensically image the device, preserving a snap shot of the file's metadata ensuring that an original version is available.

A. What Is a Forensic Image?

Forensic imaging creates a digital copy of a hard drive or device. Unlike the "duplicate" or "copy" functions of a computer, the imaged drives contain all the original metadata and the "deleted" information in the empty space of the hard drive.[211] The key to any forensic image is preserving all of the information, data, metadata, and empty space. The principles of a forensic image are similar to creating a complete back-up of a computer or a server. The difference is, the image is not created to restore a computer in the event of hardware failure but rather to preserve a "snap shot" of exactly what information was on a computer when the image was created. There are a number of companies that specialize in creating images of computers[212] and companies that create hardware to image hard drives. [213]

The computer information at issue in *Zubulake* was an image of a computer system. Unlike the forensic images addressed here, the *Zubulake* server tape could not be read or reviewed without recopying the entire tape to the server, a process that could take hours or days.[214] Had a forensic image of the drive like the ones proposed here been available, the relevant emails, and corresponding metadata if necessary, could have been retrieved quickly and easily even though the emails were deleted from the server itself.

A forensically sound image of a computer will record every single piece of information on the drive. This includes the "deleted" files that remain in

[211] Craig Ball, *Computer Forensics for Lawyers Who Can't Change the Clock on Their VCRs*, *in* SIX ON FORENSICS: SIX ARTICLES ON COMPUTER FORENSICS FOR LAWYERS 43 (2005), http://www.craigball.com/_OFFLINE/cf_vcr.pdf.

[212] *E.g.*, FORENSIC DIGITAL IMAGING INC., http://www.fdiflorida.com/ (last visited May 2, 2016); IMAGING FORENSICS, http://www.imagingforensics.com/ (last visited May 2, 2016).

[213] *E.g.*, INTELLIGENT COMPUT. SYS., http://www.ics-iq.com/Forensic-Acquisition-Lab-on-the-Road-s/35.htm (last visited May 2, 2016).

[214] *See* Zubulake I, 217 F.R.D. 309, 313 (S.D.N.Y. 2003).

the empty space of a computer.[215] The process of making the image does not modify any of the information on the original computer; it is preserved just as if the image was never created. Thus, lawyers and clients are assured that the original metadata and computer information are preserved should it later be required in the course of litigation.

The cost of forensically imaging a drive is, admittedly, not insignificant. Imaging a hard drive costs between $300 and $1,250 depending on the drive size and other factors.[216] Of course, complicated litigation may require imaging dozens of drives, which in turn requires maintaining a large computer system to store the images and secure them from hacking and unauthorized access. However, creating more demand for forensic imaging services will help encourage innovation in the marketplace, which will help lower costs over time. Larger firms may be able to afford their own in-house imaging services to further reduce costs.

The cost may be prohibitive for some legal matters that do not involve large dollar amounts or lack the complexity to be worth the client's or lawyer's time and expense to image the drive. Other cases that may not require a forensic image to adequately preserve information; the parties can stipulate to the authenticity of documents or the contents of metadata. Simpler litigation likely will not involve the lawyer taking possession of clients' devices or hard drives. In simple cases where metadata will not offer useful information or neither party wants to undergo a detailed metadata analysis, the parties can stipulate to the authenticity of electronic information or other relevant facts that metadata may show. In cases where only a handful of documents are at issue, especially when the documents can be transmitted via e-mail, a secure website, or other means that does not involve the swapping of computer hardware, then there is nothing to image. In those cases, backing up the metadata can be as easy as working off a duplicate of the file. In the case of documents received by e-mail, it is sufficient to preserve the e-mails containing the documents and insuring they are not deleted by automatic archival software. At most, firms need to make a one-time (or occasional)

[215] Ball, *supra* note 211 ("A 'forensically-sound' duplicate of a drive is, first and foremost, one created by a method which does not, in any way, alter any data on the drive being duplicated. Second, a forensically-sound duplicate must contain a copy of every bit, byte and sector of the source drive, including unallocated 'empty' space and slack space, precisely as such data appears on the source drive relative to the other data on the drive. Finally, a forensically-sound duplicate will not contain any data (except known filler characters) other than which was copied from the source drive.").

[216] *See, e.g.*, *Hard Drive Imaging Fees*, E-DISCOVERY INC., http://www.ediscoveryinc.com/services/computer-forensics/hard-drive-imaging-fees/ (last visited May 16, 2016).

purchase of archival software to preserve an archive of all the relevant information.

Though this does require firms to invest in computer storage sufficient to store the images, computer storage costs have declined dramatically over the years.[217] Moreover, lawyers would not have to preserve the images indefinitely; they would only need to be preserved for the same amount of time that physical documents are normally retained by the firm.

B. Why Create a Forensic Image?

Spoliation of evidence is a significant risk when litigation involves voluminous electronic information. Computer files may be exchanged between the client and the lawyer, the lawyer and expert witnesses, the lawyers of both (or in complex litigation many) clients, and the lawyers and the court, and, of course, between lawyers in the firm working on the same case. All of these exchanges risk altering metadata and precluding successful authentication of certain files. It is easy to place the responsibility to preserve metadata on the clients, and in most cases that is sufficient. However, clients sometimes make mistakes, and computer systems are not infallible or static; software updates, computer upgrades, new software, and routine maintenance, can sometimes change documents or remove key information.[218]

As the advisory opinions from state ethics boards indicate, clients sometimes lose their files and want to retrieve them from the lawyers.[219] In the event that this happens during litigation or the files are at issue in subsequent litigation, the lawyers' version of a file may be the only one available. If those files' metadata are altered, the opposing party can

[217] Matthew Komorowski, *A History of Storage Cost (Update)*, MKOMO, http://www.mkomo.com/cost-per-gigabyte-update (last updated Mar. 9, 2014).

[218] For example: in 2003, Microsoft released a new version of Word; the new version no longer preserved a running list of all the people who accessed a particular file and deleted any list in files created by the previous version of Word whenever it was accessed by the updated software. As a result, key metadata about the files were deleted even though the data appeared unaltered. Tanner, *supra* note 43, at 4.

[219] Ala. State Bar Ethics Op., Formal Op. 2010-02 (2010), https://www.alabar.org/resources/office-of-general-counsel/formal-opinions/2010-02/ (discussing retention, storage, ownership, production and destruction of client files); State Bar of Ariz. Ethics Opinions, Formal Op. 09-04 (2009), http://www.azbar.org/Ethics/EthicsOpinions/ViewEthicsOpinion?id=704 (discussing confidentiality; maintaining client files; electronic storage; internet); State Bar of Nev. Standing Comm. on Ethics and Prof'l Responsibility, Formal Op. No. 23 (2006), http://ftp.documation.com/references/ABA10a/PDfs/3_12.pdf.

challenge the authenticity of documents or files. In that event, it becomes a battle of the computer data against the word of the parties involved, who no doubt are acting at least in part out of self-interest. As the Third Circuit noted, it is and will continue to become more difficult to determine when the original documents need to be preserved.[220]

Given the potential to inadvertently modify metadata, the best practice for lawyers when they receive clients' files is to preserve a copy with all of the metadata and invisible files. Thus, when taking possession of a client's computer, hard drive, or flash drive in the course of litigation, the lawyer should create a forensic image of the drive *before* opening, modifying, redacting, or reorganizing any of the documents.[221] Even if metadata are not at issue at the start of litigation, there is a chance that it will become relevant as discovery progresses. Thus, it is better to be cautious and preserve computer files in their original form with the metadata and work off of duplicates for any redactions or modifications. While it does take time to image a drive, it does not necessarily require a lawyer to supervise the process. A trained information technology person could easily image the drive and provide the files to the lawyer when the imaging is complete.

Organizations and companies may be under a limited duty to preserve metadata; however, the Sedona Conference has observed that companies are generally not required to preserve metadata unless it is on notice that metadata are at issue in a dispute.[222] The Conference does not address the lawyer's duty to preserve documents received in preparing responses to discovery requests. The ABA has advised, based on a California opinion, that lawyers retain the originals of any documents they receive.[223] However, the ABA's opinion is concerned with physical files and makes no mention of preserving original electronic information. The forthcoming revisions to the

[220] Bull v. United Parcel Serv., Inc., 665 F.3d 68, 78–79 n.12 (3d Cir. 2012).

[221] Forensic images can also be referred to as clones, bit stream copies, ghosts, or a mirror; so long as all the data is preserved, it does not matter what it is called. Ball, *supra* note 211.

[222] THE SEDONA GUIDELINES, *supra* note 169, at 28, 30 ("Absent a legal requirement to the contrary, organizations are not required to preserve metadata, but may find it useful to do so in some instances. . . . [I]f in the ordinary course of business an organization migrates electronic versions with associated metadata to other versions without retaining that metadata, the organization should consider if and how it would preserve electronic versions including metadata if it has actual notice (by court order or otherwise) that the metadata is [sic] material and needs to be preserved.").

[223] Ctr. for Prof'l Responsibility, *Materials on Client File Retention*, ABA http://www.americanbar.org/groups/professional_responsibility/services/ethicsearch/materials_on_client_file_retention.html (last visited May 16, 2016) ("As to the question of what should be retained and when items may be destroyed, see California Opinion 2001-157 (undated) that states that a lawyer must retain original papers and property received from a former client, including estate planning documents, according to the law of deposits and the Probate Code.").

Federal Rules of Civil Procedure are equally silent about preserving metadata, though the consequences for negligent spoliation will be more consistent across federal jurisdictions.

Electronic files are more dynamic and contain a host of important information beyond the content displayed on screen. The duty to preserve metadata has not *yet* been defined by the rules of civil procedure or a court; however, neither had the duty to preserve ESI been clearly defined prior to *Zubulake*.[224] It is only a matter of time before a court is forced to draw the line. For now, the best practice is to forensically image any clients' hard drives or devices a lawyer takes possession of during litigation.[225]

To be clear, I do not suggest that hard drives should be analyzed to search for deleted files to extract metadata in *every* case. Indeed, depending on the litigation, the parties can stipulate what will be discoverable or relevant or even when files were created and minimize concern over metadata spoliation. However, the raw computer information should be preserved in case such analyses are necessary later in the discovery process. The best practice of attorneys working with files and devices attained from their clients is to create a forensic image of the drive and preserve a copy of all of the information on the drive. If questions later arise about metadata or deleted files, the forensic images will be readily available for a more detailed analysis. If a hearing is required to determine whether the lawyers inadvertently altered metadata (or any ESI), the forensic images should readily answer that question. In low dollar cases where the money at issue makes it too expensive to create a forensic image and still proceed with the litigation, the practitioner should discuss these issues with the client and explain that while the best practice is to create an image of the file, it is not yet routine and may be costly. In the course of informing the client, the practitioner should also warn the client that failing to preserve the metadata could potentially lead to sanctions of various degrees.[226] However, practitioners should preserve the information whenever it is feasible.

In high-stakes or complex cases where thousands of documents are exchanged between clients and their attorneys, the parties can agree to share

[224] Li, *supra* note 160.

[225] Preserving that much digital information creates cyber security issues for firms. However, those concerns should already be on attorneys' minds whenever they handle sensitive client files. Such concerns and implications are outside the scope of this article. For a snapshot of how sophisticated clients are pressuring law firms on cyber security, see Mathew Goldstein, *Law Firms Are Pressed on Security for Data*, N.Y. TIMES (Mar. 26, 2014, 7:00 PM), http://dealbook.nytimes.com/2014/03/26/law-firms-scrutinized-as-hacking-increases.

[226] If a client wants the practitioner to create a forensic image after this conversation, the practitioner should inform the client of the potential cost.

the burden of metadata preservation. Indeed, some clients may prefer to maintain the forensic images themselves. If the lawyer and client agree to an arrangement that shifts some or all of the burden of forensically imaging to the client, the lawyer would still need to ensure that the image is sufficiently sound to preserve the information and preserved for the duration of the litigation. When the client maintains possession of the images, the lawyers' duty to ensure preservation is roughly equivalent to what it is now under the Federal Rules of Civil Procedure and as outlined in the *Zubulake* decisions.

Rather than suggesting that the duty to preserve ESI should shift from the client to the lawyer, I propose a new duty that starts when a lawyer takes possession of client documents during litigation. The client is still under a duty to preserve the information, and the lawyer is still under some duty to ensure that a client follows litigation hold orders.[227] Further, a proposal that lawyers take possession of every client computer or hard drive would be far too burdensome on the lawyer and costly to the client. The lawyers' and clients' respective jobs in this regard have been carefully established, and those roles need not change. Indeed, clients should retain possession of their files and comply with litigation hold orders, while lawyers should focus on making sure relevant and potentially relevant files are not destroyed inadvertently (or intentionally) in the course of ordinary business. Even when the client retains the originals (or copies) of the documents, the best practice should be that the lawyer backs-up the files receives from the client and ensure that the lawyer's action reviewing, redacting, and imaging[228] the files do not alter or delete all of the metadata in the files. If metadata are not at issue at the start of litigation but becomes relevant later in discovery, the information is already preserved and ready for analysis or extraction.

IV. CONCLUSION

The world has changed. The computer revolution has forever altered the landscape of our broader society and has also forever altered litigation and discovery. The lawyer's role still involves carefully scrutinizing files for relevance, privilege, and smoking guns, but now incorporates navigating the complexity of dynamic computer files that are not as static as the accustomed paper files. Questions about what to do with metadata are not going away and

[227] Hold orders are letters transmitted to clients, particularly business, informing them of filed or potential litigation and advising them to preserve all potentially relevant evidence.

[228] In this context, I use "imaging" to describe the process of stripping a file of its metadata and hidden data (pseudo-metadata) by exporting or saving it as a PDF or TIFF image. *See supra* notes 68–70 and accompanying text.

will probably only intensify in the coming years, and lawyers should more aggressively preserve this evidence. When a lawyer receives a hard drive or device from a client, the first thing the lawyer should do is forensically image it or take other reasonable steps to preserve the metadata. In more complex, high-stakes cases, lawyers should preserve the metadata received from their clients. Until the courts give firm guidance on metadata and its relevancy, the best practice is to preserve the information and ensure that if the client does not properly preserve the metadata or if it later becomes relevant, it is readily available and no action by the lawyer has altered or damaged the metadata. At a minimum, lawyers should appreciate that the ESI disputes of the last few years are a shadow of the complex world of computer information that will permeate litigation for the foreseeable future.