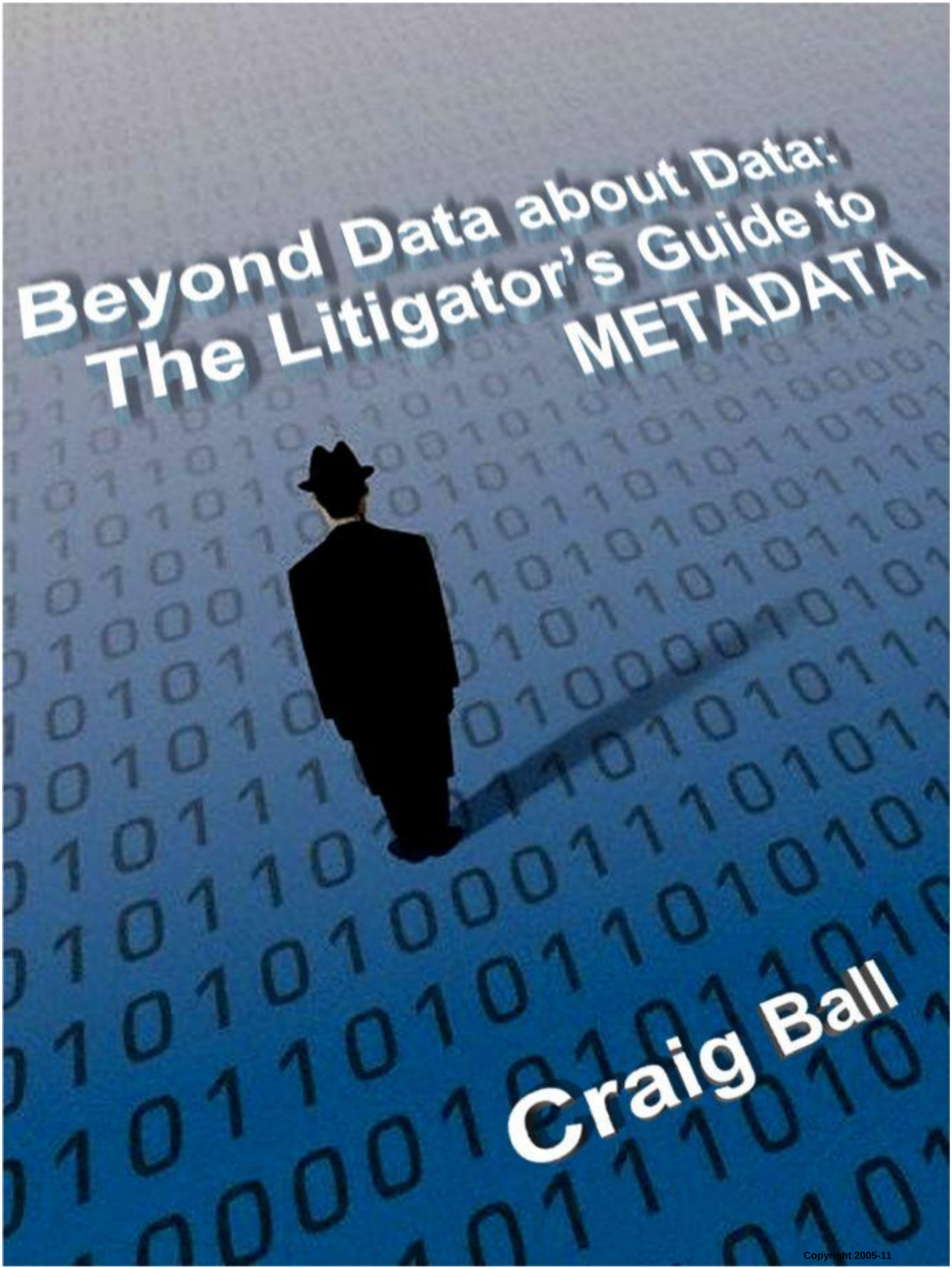


# Beyond Data about Data: The Litigator's Guide to METADATA

Craig Ball

# Beyond Data about Data: The Litigator's Guide to Metadata

By Craig Ball

In the old joke, a balloonist descends through the fog to get directions. "Where am I?" she calls out to a man on the ground, who answers, "You're in a yellow hot air balloon about sixty-seven feet above the ground." The frustrated balloonist replies, "Thanks for nothing, Counselor." Taken aback, the man on the ground asks, "How did you know I'm a lawyer?" "Simple," says the balloonist, "your answer was 100% accurate and totally useless."

If you ask a tech-savvy lawyer, "What's metadata?" there's a good chance you'll hear, "Metadata is data about data." Another answer that's 100% accurate and totally useless.

It's time to move past "data about data" and embrace more useful ways to describe metadata—ways that enable counsel to rationally assess relevance and burden attendant to metadata. Metadata is probably the most misunderstood topic in electronic discovery today. All too frequently, lawyers for requesting parties demand discovery of "the metadata" without specifying what metadata is sought, and lawyers for producing parties fail to act to preserve metadata of genuine value and relevance.

**It's time to get past defining metadata as "data about data."**

**It's Information *and* Evidence**

Metadata is information that helps us use and make sense of other information. More particularly, metadata is evidence, typically stored electronically, that describes the characteristics, origins, usage, structure, alteration and validity of other electronic evidence. Many instances of metadata in many forms occur in many locations within and without digital files. Some is supplied by the user, but most metadata is generated by systems and software. Some is crucial evidence and some is just digital clutter. Appreciating the difference--knowing what metadata exists and understanding its evidentiary significance--are skills essential to electronic discovery.

**Metadata is Evidence?**

If evidence is anything that tends to prove or refute an assertion as fact, then clearly metadata is evidence. Metadata sheds light on the origins, context, authenticity, reliability and distribution of electronic evidence, as well as provides clues to human behavior. It's the electronic equivalent of DNA, ballistics and fingerprint evidence, with a comparable power to exonerate and incriminate.

In ***Williams v. Sprint/United Mgmt Co.***, 230 F.R.D. 640 (D. Kan. 2005), the federal court ruled:

> [W]hen 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.

Within the realm of metadata lies discoverable evidence that litigants are obliged to preserve and produce. There's as much or more metadata extant as there is information and, like information, you don't deal with every bit of it. You *choose* wisely.

A lawyer's ability to advise a client about how to find, preserve and produce metadata, or to object to its production and discuss or forge agreements about metadata, hinges upon how well he or she understands metadata.

## It's Just Ones and Zeroes

Understanding metadata and its importance in e-discovery begins with awareness that electronic data is, fundamentally, just a series of ones and zeroes. Though you've surely heard that before, you may not have considered the implications of information being expressed so severely. There are no *words*. There are no spaces or punctuation. *There is no delineation of any kind*.

How, then, do computers convert this unbroken sequence of ones and zeroes into information that makes sense to human beings? There has to be some *key*, some *coherent structure* imposed to divine their meaning. But where does it come from? We can't derive it *from* the data if we can't first make sense *of* the data.

## It's Encoded

Consider that written English conveys all information using fifty-two upper- and lowercase letters of the alphabet, ten numerical digits (0-9), some punctuation marks and a few formatting conventions, like spaces, line feeds, pages, etc. You can think of these collectively as a seventy- or eighty-signal "code." In turn, much of the same information could be communicated or stored in Morse code, where a three-signal code composed of dot, dash and pause serves as the entire "alphabet."

We've all seen movies where a tapping sound is heard and someone says, "Listen! It's Morse code!" Suddenly, the tapping is an encoded *message* because someone has furnished metadata ("It's Morse code!") *about* the data (tap, tap, pause, tap). Likewise, all those ones and zeroes on a computer only make sense when other ones and zeroes—the metadata—communicate the framework for parsing and interpreting the data stream.

**All those ones and zeroes on a computer only make *sense* when other ones and zeroes—the metadata—communicate the framework for parsing and interpreting the data**

So, we need data *about* the data. We need information that tells us the data's encoding scheme. We need to know when information with one purpose ends and different information begins. And we need to know the context, purpose, timeliness and origin of information for it to help us. That's ***metadata***.

## The Metadata Continuum

Sometimes metadata is elemental, like the contents of a computer's master file table detailing where the sequences of one and zeroes for particular files begin and end. This metadata is invisible to a user without special tools called hex editors capable of peering through the walls of the Windows interface into the utilitarian plumbing of the operating system. Without file location metadata, every time a user tries to access a file or program, the operating system would have to examine every one and zero to find it. It'd be like looking for someone by knocking on every door in town!

At other times, metadata supports enhanced functionality not essential to the operation of the system. The metadata that tracks a file's name or the dates a file was created, last accessed and last modified may only occasionally be probative of an issue in the case, but that information *always* makes it easier to locate, sort and segregate files.

Metadata may be instrumental to the intelligibility of information, helping us use and make sense of it. "Sunny and 70 degrees" isn't a very useful forecast without metadata indicating *when* and *where* it's predicted to be the weather. Similarly, fully understanding information on a website or within a database, a collaborative environment like Microsoft's SharePoint or a social network like Facebook depends on metadata that defines its location, origin, timing and structure. It's even common for computerized information to comprise *more* metadata than data, in the same way that making sense of the two data points "sunny" and "70 degrees" requires *three* metadata points: location, date and time of day.

### There's No Such Thing as "The Metadata"

As we move up the evolutionary ladder for metadata, some is recorded just in case it's needed to support a specialized task for the operating system or an application. Standard system metadata fields like "Camera Model" or "Copyright" may seem an utter backwater to a lawyer concerned with spreadsheets and word processed documents, but if the issue is the authenticity of a photograph or pirated music, these fields can make or break the case. ***It's all about relevance and utility.***

The point is, there's really no such thing as "the metadata" for a file or document. Instead, there's a continuum of metadata that enlightens many aspects of ESI. The metadata that matters depends upon the issues presented in the case; consequently, the metadata preserved for litigation should reasonably reflect the issues that were—or which should have been—reasonably anticipated.

### Up by the Bootstraps

When you push the power button on your computer, you trigger an extraordinary expedited education that takes the machine from insensible illiterate to worldly savant in a matter of seconds. The process starts with a snippet of data on a chip called the **ROM BIOS** storing just enough information in its **R**ead **O**nly **M**emory to grope around for the **B**asic **I**nput and **O**utput **S**ystem devices like the keyboard, screen and hard drive. It also holds the metadata needed to permit the computer to begin loading ones and zeroes from storage and to make just enough sense of their meaning to allow more metadata to load from the disk, in turn enabling the computer to access more data and, in this widening gyre, "teach" itself to be a modern, capable computer.

This rapid, self-sustaining self-education is as magical as if you hoisted yourself into the air by pulling on the straps of your boots, which is truly why it's called "bootstrapping" or just "booting" a computer.

### File Systems

So now that our computer's taught itself to read, it needs a library. Most of those ones and zeroes on the hard drive are files that, like books, are written, read, revised and referenced. Computers use file systems to keep track of files just as libraries once used card catalogues and the Dewey Decimal system to track books.

Imagine you own a thousand books without covers that you stored on one very long shelf. You also own a robot that can't read, but Robby can count books very accurately. How would you instruct Robby to get a particular book?

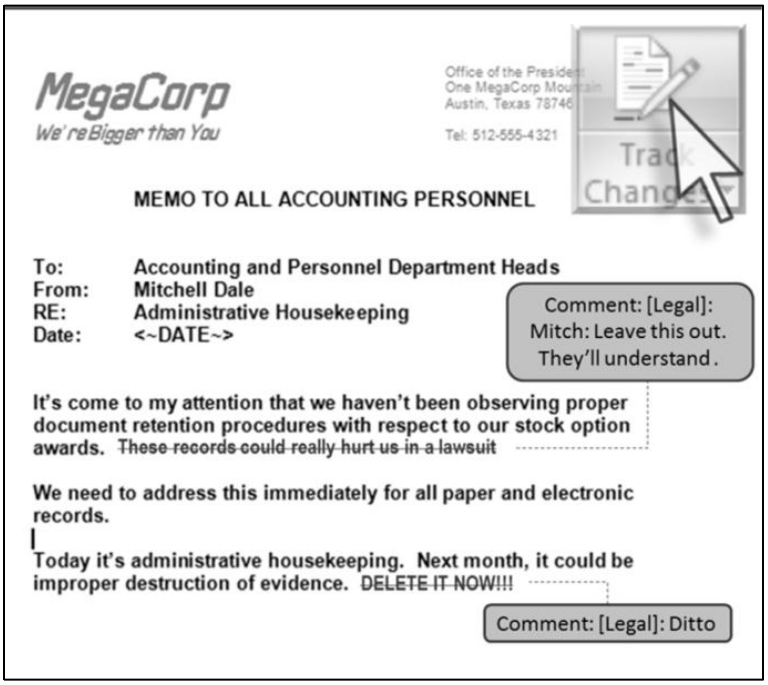
If you know the order in which the books are stored, you'd say, "Robby, bring me the 412th book." If it was a set of encyclopedias, you'd add: "…and the next 23 books." The books don't "know" where they're shelved. Each book's location is metadata *about* the book.

Locating something by specifying that it's so many units from a particular point is called *relative addressing*. The number of units the destination is set off from the specified point is called the *offset*. Computers use offset values to indicate the locations of files on storage devices as well as to locate particular information within files.

Computers use various units to store and track information, so offsets aren't always expressed in the same units. A "bit" stores a one or zero, eight bits is a "byte," (sufficient to hold a letter in the Latin alphabet), 512 bytes is a *sector or block* (see **Appendix A**) and (typically) eight contiguous sectors or blocks is a *cluster*. The cluster is the most common unit of logical storage, and modern computers tend to store files in as many of these 4,096-byte clusters, or "data baskets," as needed. Offset values are couched in bytes when specifying the location of information within files and as sectors when specifying the location of files on storage media.

**Metadata Mix-up: Application Metadata**

To the extent lawyers have heard of metadata at all, it's likely in the context of just one species of metadata called **application metadata** with the fearsome potential to reveal confidential or privileged information embedded within electronic documents. Computer programs or "applications" store work product in files "native" to them, meaning that the data is structured and encoded to support the application. As these applications added features--like the ability to undo changes in or collaborate on a document--the native files used to store documents had to retain those changes and collaborations.



MegaCorp
We're Bigger than You

Office of the Preside
One MegaCorp Mou
Austin, Texas 78746

Tel: 512-555-4321

MEMO TO ALL ACCOUNTING PERSONNEL

To: Accounting and Personnel Department Heads
From: Mitchell Dale
RE: Administrative Housekeeping
Date: <~DATE~>

Comment: [Legal]: Mitch: Leave this out. They'll understand.

It's come to my attention that we haven't been observing proper document retention procedures with respect to our stock option awards. ~~These records could really hurt us in a lawsuit~~

We need to address this immediately for all paper and electronic records.

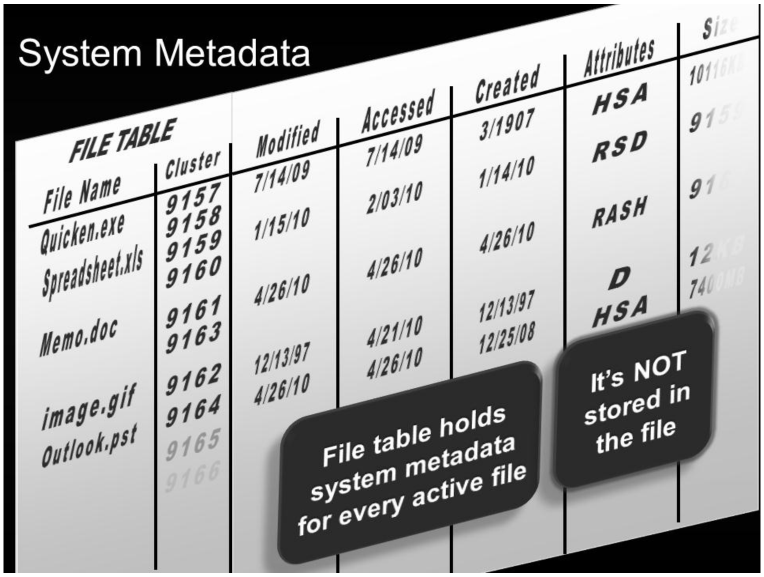
Today it's administrative housekeeping. Next month, it could be improper destruction of evidence. ~~DELETE IT NOW!!!~~

Comment: [Legal]: Ditto

An oft-cited culprit is Microsoft Word, and a cottage industry has grown up offering utilities to strip embedded information, like comments and tracked changes, from Word documents. Because of its potential to embarrass lawyers or compromise privilege, metadata has acquired an unsavory reputation

amongst the bar. But metadata is much more than simply the embedded *application* metadata that affords those who know how to find it the ability to dredge up a document's secrets.
By design, application metadata is embedded in the file it describes and moves with the file when you copy it. However, not all metadata is embedded (for the same reason that cards in a library card catalog aren't stored between the pages of the books). You have to know where the information resides to reach it.

**System Metadata**
Unlike books, computer files aren't neatly bound tomes with names embossed on spines and covers. Often, files don't internally reflect the name they've been given or other information about their location, history or ownership. The information about the file which is *not* embedded within the file it describes but is stored externally is its **system metadata**. The computer's file management system uses *system* metadata to track file locations and store demographics about each file's name, size, creation, modification and usage.



System metadata is crucial to electronic discovery because so much of our ability to identify, find, sort and cull information depends on its system metadata values. For example, system metadata helps identify the custodians of files, when files were created or altered and the folders in which they were stored. System metadata stores much of the *who, when, where* and *how* of electronic evidence.

Every computer employs one or more databases to keep track of system metadata. In computers running the Windows operating system, the principal "card catalog" tracking system metadata is called the Master File Table or "MFT." In the predecessor DOS operating system, it was called the File Allocation Table or "FAT." The more sophisticated and secure the operating system, the greater the richness and complexity of the system metadata in the file table.

**Windows Shell Items**
In the Windows world, Microsoft calls any single piece of content, such as a file, folder, email message or contact, a "**Shell item.**" Any individual piece of metadata associated with a Shell item is called a "**property**" of the item. Windows tracks 284 distinct metadata properties of Shell items in 28 property categories, as set out fully in **Appendix B**. To see the list of Shell item properties on your own Windows system, right click on the column names in any folder view and select "More…." Examining just a handful of these in four key categories reveals metadata of great potential evidentiary value existing within and without files, messages and photos:

| Category | Properties | |
|---|---|---|
| Document | ClientID<br>Contributor<br>DateCreated<br>DatePrinted<br>DateSaved<br>DocumentID | LastAuthor<br>RevisionNumber<br>Template<br>TotalEditingTime<br>Version |
| Message | AttachmentContents<br>AttachmentNames<br>BccAddress<br>BccName<br>CcAddress<br>CcName<br>ConversationID<br>ConversationIndex<br>DateReceived<br>DateSent<br>Flags | FromAddress<br>FromName<br>HasAttachments<br>IsFwdOrReply<br>SenderAddress<br>SenderName<br>Store<br>ToAddress<br>ToDoFlags<br>ToDoTitle<br>ToName |
| Photo | CameraManufacturer<br>CameraModel | CameraSerialNumber<br>DateTaken |
| System | ApplicationName<br>Author<br>Comment<br>Company<br>ComputerName<br>ContainedItems<br>ContentType<br>DateAccessed<br>DateAcquired<br>DateArchived<br>DateCompleted<br>DateCreated<br>DateImported<br>DateModified<br>DueDate<br>EndDate<br>FileAttributes<br>FileCount<br>FileDescription<br>FileExtension<br>FileName<br>IsShared | ItemAuthors<br>ItemDate<br>ItemFolderNameDisplay<br>ItemFolderPathDisplay<br>ItemName<br>OriginalFileName<br>OwnerSID<br>Project<br>Sensitivity<br>SensitivityText<br>SharedWith<br>Size<br>Status<br>Subject<br>Title<br>FileOwner<br>FlagStatus<br>FullText<br>IsAttachment<br>IsDeleted<br>IsEncrypted |

## Much More Metadata

The 284 Windows Shell item properties are by no means an exhaustive list of metadata. Software applications deploy their own complements of metadata geared to supporting features unique to each application. E-mail software, word processing applications and spreadsheet, database, web browser and presentation software collectively employ hundreds of additional fields of metadata.

For example, digital photographs can carry dozens of embedded fields of metadata detailing information about the date and time the photo was taken, the camera, settings, exposure, lighting, even precise geolocation data. Many are surprised to learn that photos taken with cell phones having GPS capabilities (like the Apple iPhone) contain detailed information about where the photo was taken *to a precision of about ten meters*.

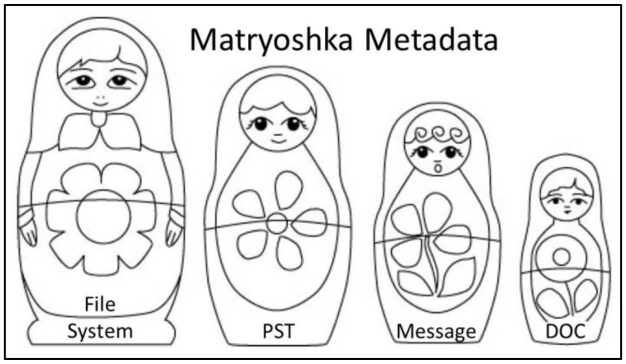
**Photos taken with cell phones having GPS capabilities…contain detailed information about where the photo was taken.**

The popular Microsoft Outlook e-mail client application provides for more than 180 standard application metadata fields which users may select to customize their view. These are set out fully in **Appendix C.**

But even this broad swath of metadata is still only *part* of the probative information about information recorded by computers. Within the Master File Table and index records used by Windows to track all files, still more attributes are encoded in hexadecimal notation. In fact, an ironic aspect of Windows is that the record used to track information about a file may be larger than the file itself! Stored within the hives of the System Registry—the "Big Brother" database that tracks attributes covering almost any aspect of the system—are thousands upon thousands of attribute values called "registry keys." Other records and logs track network activity and journal virtually every action.

### Matryoshka Metadata

Matryoshka are carved, cylindrical Russian dolls that nest inside one another. It's helpful to think of computer data the same way. If the evidence of interest is a Word document attached to an e-mail, the document has its usual complement of application metadata that moves with the file; but, as it nests within an e-mail message, its "system" metadata is only that which is contained within the transporting message. The transporting message, in turn, carries its own metadata concerning transit, addressing, structure, encoding and the like. The message is managed by Outlook, which maintains a rich complement of metadata about the message and about its own configuration. As configured, Outlook may store all messages and application metadata in a container file called Outlook.PST. This container file exists within a file system of a computer that stores system metadata about the container file, such as where the file is stored, under whose user account, when it was last modified, its size, name, associated application and so on.



Within this Matryoshka maelstrom of metadata, some information is readily accessible and comprehensible while other data is so Byzantine and cryptic as to cause even highly skilled computer forensic examiners to scratch their heads.

### Forms of Metadata

Now that your head is spinning from all the types, purposes and sources of metadata, let's add another complexity concern: the *form* of the metadata. Metadata aren't presented the same way from field to field or application to application.  For example, some of the standard metadata fields for Outlook e-mail are simply bit flags signifying “true” or “false” for “Attachment,” “Do Not Auto Archive,” “Read” or “Receipt Requested.”  Some fields reference different units, e.g., “Size” references bytes, where “Retrieval Time” references minutes.  Several fields even use the *same* value to mean *different* things, e.g., a value of “1” signifies “Completed” for “Flag Status,” but denotes “Normal for “Importance,” “Personal” for “Sensitivity” and “Delivered” for “Tracking Status.”

The form of metadata is a key consideration when deciding how to preserve and produce the information.  Not everyone would appreciate a response like, “for this message, item type 0x0029 with value type 0x000b was set to 0x00,” when the question posed was whether the sender sought a read receipt.  Because some metadata items are simply bit flags or numeric values and make sense only as they trigger an action or indication in the native application, preserving metadata can entail more than just telling opposing counsel, “we will grab it and give it to you.”

It’s not that locating and interpreting any particular item is difficult, but you have to know whether your firm, client or service provider has the tools and employs a methodology that makes it easy. That’s why it’s crucial to know what metadata is routinely collected and amenable to production before making commitments to opposing counsel or the court.  Any e-discovery vendor you employ should be able to readily identify the system and application metadata values they routinely collect and process for production.  Any still-existing metadata value can be collected and processed; but, some require specialized tools and software, custom programming or changes to established workflows.

**Relevance and Utility**

How much of this metadata is relevant and discoverable?  Would I be any kind of lawyer if I didn’t answer, “It depends?”  In truth, it *does* depend upon what issues the data bears upon, its utility and the cost and burden of preservation and review.

Metadata is unlike almost any other evidence in that its import in discovery may flow from its probative value (relevance as evidence), its utility (functionally abetting the searching, sorting and interpretation of ESI) or both.  If the origin, use, distribution, destruction or integrity of electronic evidence is at issue, the relevant “digital DNA” of metadata is essential, probative evidence that needs to be preserved and produced.  Likewise, if the metadata materially facilitates the searching sorting and management of electronic evidence, it should be preserved and produced for its utility.

Does this then mean that every computer system and data device in every case must be forensically imaged and analyzed by experts?  Absolutely not!  *Once we understand what metadata exists and what it signifies, a continuum of reasonableness will inform our actions*.  A competent police officer making a traffic stop collects relevant information, such as, e.g., the driver’s name, address, vehicle license number, driver’s license number and date, time and location of offense.  We wouldn’t expect the traffic cop to collect a bite mark impression, DNA sample or shoe print from the driver.  But, make it a murder case and the calculus changes.

Addressing just the utility aspect of metadata in the context of forms of production, The Sedona Conference guideline states:

> Absent party agreement or court order specifying the form or forms of production, production should be made in the form or forms in which the information is ordinarily maintained or in a reasonably usable form, **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 where appropriate or necessary in light of the nature of the information and the needs of the case.**

*The Sedona Principles Addressing Electronic Document Production, Second Edition (June, 2007),* Principle 12 (emphasis added).

The crucial factors are burden and cost balanced against utility and relevance. The goal should be a level playing field between the parties in terms of their ability to see and use relevant electronic evidence.

**Periodically re-assess the adequacy of preservation and production and act to meet changed circumstances**.

So where do we draw the line? Begin by recognizing that the advent of electronic evidence hasn't changed the fundamental dynamics of discovery: *Litigants are entitled to discover relevant, non-privileged information, and relevance depends on the issues before the court.* Relevance assessments aren't static, but change as new evidence emerges and new issues arise. Metadata irrelevant at the start of a case may become decisive when, e.g., allegations of data tampering or spoliation emerge. Parties must periodically re-assess the adequacy of preservation and production of metadata and act to meet changed circumstances.

**Metadata Musts**

There are easily accessible, frequently valuable metadata that, like the information collected by the traffic cop, we should expect to routinely preserve. Examples of essential system metadata fields for any file produced are:

- **Custodian;**
- **Source Device;**
- **Originating Path** (File path of the file as it resided in its original environment);
- **Filename** (including extension);
- **Last Modified Date;** *and*
- **Last Modified Time**.

Any party producing or receiving ESI should be able to state something akin to, "This spreadsheet named *Cash Forecast.xls* came from the My Documents folder on Sarah Smith's Dell laptop and was last modified on January 16, 2011 at 2:07 PM CST."

One more metadata "must" for time and date information is the UTC time zone offset applicable to each time value (unless all times have been normalized; that is, processed to a common time zone). UTC stands for both for Temps Universel Coordonné and Coordinated Universal Time. It's a fraction of a second off the better known Greenwich Mean Time (GMT) and identical to Zulu time in military and aviation circles. Why UTC instead of TUC or CUT? It's a diplomatic compromise, for neither French nor English speakers were willing to concede the acronym. Because time values may be expressed with reference to local time zones and variable daylight savings time rules, you need to know the UTC offset for each item.

Application metadata is, by definition, embedded within native files; so, native production of ESI obviates the need to selectively preserve or produce application metadata. It’s in there. But when ESI is converted to other forms, the producing party must assess what metadata will be lost or corrupted by conversion and identify, preserve and extract relevant or utile application metadata fields for production.

For e-mail messages, this is a fairly straightforward process, notwithstanding the dozens of metadata values that may be introduced by e-mail client and server applications. The metadata “musts” for e-mail messages are:

- **Custodian** –Owner of the mail container file or account collected;
- **To** –Addressee(s) of the message;
- **From** –The e-mail address of the person sending the message;
- **CC** –Person(s) copied on the message;
- **BCC** –Person(s) blind copied on the message;
- **Date Sent** –Date the message was sent;
- **Time Sent** –Time the message was sent;
- **Subject** –Subject line of the message;
- **Date Received** –Date the message was received;
- **Time Received** –Time the message was received;
- **Attachments** –Name(s) or other unique identifier(s) of attachments;
- **Mail Folder Path** –Path of the message to its folder from the root of the originating mail account;
- **Message ID** –Microsoft Outlook or similar unique message identifier; and
- **Text** –The extracted text of the message body.

E-mail messages that traverse the Internet contain so-called header data detailing the routing and other information about message transit and delivery. Whether header data should be preserved and produced depends upon the reasonable anticipation that questions concerning authenticity, receipt or timing of messages will arise.

The metadata essentials must further include metadata values generated by the discovery and production process itself, such as Bates numbers and ranges, hash values, production paths and names, extracted or OCR text, family designations and the like.

When ESI other than e-mail is converted to non-native forms, it can be enormously difficult to preserve, produce and present relevant or necessary application metadata in ways that don’t limit its utility or intelligibility. For example, tracked changes and commentary in Microsoft Office documents may be incomprehensible without seeing them in context, i.e., superimposed on the document. By the same token, furnishing a printout or image of the document with tracked changes and comments revealed can be confusing and deprives a recipient of the ability to see the document as the user ultimately saw it. If native forms will not be produced, the most equitable approach may be to produce the document twice: once with tracked changes and comments hidden and once with them revealed.

For certain ESI, there is simply no viable alternative to native production with metadata intact. The classic example is a spreadsheet file. The loss of functionality and the confusion engendered by rows and columns that break and splay across multiple pages mandates native production. The same loss of functionality occurs with sound files (e.g., voice mail), video, animated presentations (i.e., PowerPoint) and content like databases, web content, SharePoint, social networking sites and collaborative environments where the structure and interrelationship of the information--reflected in its metadata—defines its utility and intelligibility.

***The Path to Production of Metadata***

The balance of this paper discusses steps typically taken in shepherding a metadata production effort, including:

- Gauge spoliation risks before you begin
- Identify potential forms of metadata
- Assess relevance
- Consider authentication and admissibility
- Evaluate need and methods for preservation
- Collect metadata
- Plan for privilege and production review
- Resolve production issues

- **Gauge spoliation risks before you begin**

German scientist Werner Heisenberg thrilled physicists and philosophy majors alike when he posited that the very act of observing alters the reality observed. Heisenberg's Uncertainty Principal speaks to the world of subatomic particles, but it aptly describes a daunting challenge to lawyers dealing with metadata: *When you open any document in Windows without first employing specialized hardware or software, metadata changes and prior metadata values are lost.* Altered metadata implicates not only claims of spoliation, but also severely hampers the ability to filter data chronologically. How, then, can a lawyer evaluate documents for production without *reading* them?

Begin by gauging the risk. Not every case is a crime scene, and few cases implicate issues of computer forensics. Those that do demand extra care be taken immediately to preserve a broad range of metadata evidence. Further, it may be no more difficult or costly to preserve data using forensically sound methods that reliably preserve all data and metadata.

For the ordinary case, a working knowledge of the most obvious risks and simple precautions are sufficient to protect the metadata most likely to be needed. (*See* Metadata Musts, *supra*).

1. Windows systems typically track at least three date values for files, called “MAC dates” for Last Modified, Last Accessed and Created. Of these, the Last Accessed date is the most fragile,

yet least helpful. Last accessed dates can be changed by previewing files and running virus scans. Further, last accessed dates are only infrequently updated in Windows Vista and Win7.

2. Similarly unhelpful in e-discovery is the Created date. The created date is often presumed to be the authoring date of a document, but it more accurately reflects the date the file was "created" within the file system of a particular device. So when you copy a file to new media, you're created it on the new media as of the date of copying, and the created date changes accordingly. Conversely, when you use an old file as a template to create a new document, the creation date of the template stays with the new document.

3. The date value of greatest utility in e-discovery is the Last Modified date. The last modified date of a file is not changed by copying, previewing or virus scans. It changes only when a file is opened and saved; however, it is not necessary that the user-facing content of a document be altered for the last accessed date to change. Any change—including subtle, automatic changes to application metadata--will trigger an update to the last modified date when the file is re-saved by a user.

4. Apart from corruption, application metadata does not change unless a file is opened. So, the easiest way to preserve a file's application metadata is to keep a pristine, unused copy of the file and access only working copies. By always having a path back to a pristine copy, inadvertent loss or corruption of metadata is harmless error. Calculating and preserving hash values for the pristine copies is a surefire way to demonstrate that application metadata hasn't changed

5. An approach favored by computer forensic professionals is to employ write blocking hardware or software to intercept all changes to the evidence media.

6. Finally, copies can be transferred to read only media (e.g., a CD-R or DVD-R), permitting examination without metadata corruption.

- **Identify potential forms of metadata**

To preserve metadata and assess its relevance, you have to know it exists. So for each principal file type subject to discovery, assemble a list of associated metadata of potential evidentiary or functional significance. You'll likely need to work with an expert the first time or two, but once you have a current and complete list, it will serve you in future matters. You'll want to know not only what the metadata field contains, but also its location and its significance.

For unfamiliar or proprietary applications and environments, enlist help identifying metadata from the client's IT personnel. Most importantly, *seek your opponent's input, too*. Your job is simpler when the other side is conversant in metadata and can expressly identify fields of interest. The parties may not always agree, but at least you'll know what's in dispute.

- **Assess relevance**

Are you going to preserve and produce dozens and dozens of metadata values for every document and e-mail in the case? Probably not, although you may find it easier to preserve all than selectively cull out just those values you deem relevant.

Metadata is like the weather reports from distant cities which run in the daily paper. Though only occasionally relevant, you want the information available when you need it.

Relevance is always subjective and is as fluid as the issues in the case. Case in point: two seemingly innocuous metadata fields common to Adobe Portable Document Format (PDF) files are “PDF Producer” and “PDF Version.” These are listed as “Document Properties” under the “File” menu in any copy of Adobe Acrobat. Because various programs can link to Acrobat to create PDF files, the PDF Producer field stores information concerning the source application, while the PDF Version field tracks what release of Acrobat software was used to create the PDF document. These metadata values may seem irrelevant, but consider how that perception changes if the dispute turns on a five-year-old PDF contract claimed to have been recently forged. If the metadata reveals the PDF was created using a scanner introduced to market last year and the latest release of Acrobat, that metadata supports a claim of recent fabrication. In turn, if the metadata reflects use of a very old scanner and an early release of Acrobat, the evidence bolsters the claim that the document was scanned years ago. Neither is conclusive on the issue, but both are relevant evidence needing to be preserved and produced.

Assessing relevance is another area where communication with an opponent is desirable. Often, an opponent will put relevance concerns to rest by responding, “I don’t need that.” For every opponent who demands “all the metadata,” there are many who neither know nor care about metadata.

- **Consider Authentication and Admissibility**

Absent indicia of authenticity like signature, handwriting and physical watermarks, how do we establish that electronic evidence is genuine or that a certain individual created an electronic document? Computers may be shared or unsecured and passwords lost or stolen. Software permits alteration of documents sans the telltale signs that expose paper forgeries. Once, we relied upon dates in correspondence to establish temporal relevance, but now documents may generate a new date each time they are opened, inserted by a word processor macro as a “convenience” to the user.

Where the origins and authenticity of evidence are in issue, preservation of original date and system user metadata is essential. When deciding what metadata to preserve or request, consider, *inter alia,* network access logs and journaling, evidence of other simultaneous user activity and version control data.

An important role of metadata is establishing a sound chain of custody for ESI. Through every step in e-discovery--collection, processing, review, and production—the metadata should facilitate a clear, verifiable path back to the source ESI, device and custodian.

In framing a preservation strategy, balance the burden of preservation against the likelihood of a future need for the metadata, but remember, if you act to preserve metadata for documents supporting your case, it’s hard to defend a failure to preserve metadata for items bolstering the opposition’s case. Failing to preserve metadata could deprive you of the ability to challenge the relevance or authenticity of material you produce.

- **Evaluate Need and Methods for Preservation**

Not every item of metadata is important in every case, so what factors should drive preservation? The case law, rulings of the presiding judge and regulatory obligations are paramount concerns, along with obvious issues of authenticity and relevance; but another aspect to consider is the *stability* of particular metadata. As discussed, some essential metadata fields, like Last Modified Date, change when a file is used and saved. If you don't preserve dynamic data, you lose it. Where a preservation duty has attached, by, e.g., issuance of a preservation order or operation of law, the loss of essential metadata may constitute spoliation subject to sanction.

**Fail to preserve metadata at the earliest opportunity and you may never be able to replicate what was lost.**

How, then, do you avoid spoliation occasioned by review and collection? What methods will preserve the integrity and intelligibility of metadata? Poorly executed collection efforts can corrupt metadata. When, for example, a custodian or reviewer copies responsive files to new media, prints documents or forwards e-mail, metadata is altered or lost. Consequently, metadata preservation must be addressed *before* a preservation protocol is implemented. Be certain to document what was done and why. Advising your opponents of the proposed protocol in sufficient time to allow them to object, seek court intervention or propose an alternate protocol helps to protect against belated claims of spoliation.

**Poorly executed collection efforts can corrupt metadata.**

- **Collect Metadata**

Because metadata is stored both within and without files, simply duplicating a file without capturing its system metadata may be insufficient. However, not all metadata preservation efforts demand complex and costly solutions. It's possible to tailor the method to the case in a proportional way. As feasible, record and preserve system metadata values *before* use or collection. This can be achieved using software that archives the basic system metadata values to a table, spreadsheet or CSV file. Then, if an examination results in a corruption of metadata, the original values can be ascertained. Even just archiving files ("zipping" them) may be a sufficient method to preserve associated metadata. In other cases, you'll need to employ forensic imaging or use vendors specializing in electronic discovery.

Whatever the method chosen, be careful to preserve the association between the data and its metadata. For example, if the data is the audio component of a voice mail message, it may be of little use unless correlated with the metadata detailing the date and time of the call and the identity of the voice mailbox user.

When copying file metadata, know the limitations of the environment and medium in which you're working. I learned this lesson the hard way years ago while experimenting with recordable CDs as a means to harvest files and their metadata. Each time I tried to store a file and its MAC dates (modified/accessed/created) on a CD, I found that the three *different* MAC dates derived from the hard drive would always emerge as three *identical* MAC dates when read from the CD! I learned that optical media like CD-Rs aren't formatted in the same manner as magnetic media like hard drives. Whereas the operating system formats a hard drive to store three distinct dates, CD-R media stores just one. In a sense, a CD hasn't the "slots" to store all three dates. When the CD's contents are copied back to magnetic media, the operating system re-populates the slots for the three dates with the single date found on the optical media. Thus, *using a CD in this manner serves to both corrupt and misrepresent the metadata.* Similarly, different operating systems and versions of applications maintain different metadata; so, test your processes against alteration, truncation or loss of metadata.

- **Plan for Privilege and Production Review**

The notion of reviewing metadata for privileged communications may seem odd unless you consider that application metadata potentially contains deleted content and commentary. When the time comes to review metadata for production and privilege, the risks of spoliation faced in harvest may re-appear during review. Consider:

- How will you efficiently access metadata?
- Will the metadata exist in a form you can interpret?
- Will your examination alter the metadata?
- How will you flag particular metadata for production?
- How can you redact privileged or confidential metadata?

If a vendor or in-house discovery team has extracted the metadata to a slip-sheet in an image format like TIFF or PDF, review is as simple as reading the data. However, if review will take place in native format, some metadata fields may be inaccessible, encoded or easily corrupted. If the review set is hosted online, be certain you understand which metadata fields are accessible and intelligible via the review tool and which are not. Don't just assume: *test.*

### Application Metadata and Review

Many lawyers deal with metadata in the time-honored way: *by pretending that it doesn't exist.* That is, they employ review methods that don't display application metadata, such as comments and tracked changes present in native Microsoft Office productivity documents. These lawyers review only what prints instead of *all* the information in the document. Rather than adjust their methods to the evidence, they refuse to produce ESI with its application metadata intact lest they unwittingly produce privileged or confidential content.

They defend this behavior by claiming that the burden to review application metadata for privileged or confidential content is greater than the evidentiary value of that content. To insure that requesting parties cannot access all that metadata the producing counsel ignored, producing parties instead strip away all metadata, either by printing the documents to paper or hiring a vendor to convert the ESI to static images (i.e., TIFFs). Doing so successfully removes the metadata, but wrecks the utility and searchability of most electronic evidence.

Sometimes, counsel producing TIFF image productions will undertake to reintroduce some of the stripped metadata and searchable text as ancillary productions called "**load files**." The production of document images and load files is a high-cost, low utility, error-prone approach to e-discovery; but, its biggest drawback is that it's increasingly unable to do justice to the native files it supplants. When produced as images, spreadsheets often become useless and incomprehensible. Multimedia files disappear. Any form of interactive, animated or structured information ceases to work. In general, the richer the information in the evidence, the less likely it is to survive production in TIFF.

Despite these shortcomings, lawyers cling to cumbersome TIFF productions, driving up e-discovery costs. This is troubling enough, but raises a disturbing question: Why does any lawyer assume he or she is free to unilaterally suppress--without review or proffer of a privilege log—integral parts of discoverable evidence? Stripping away or ignoring metadata that's an integral part of the evidence seems little different from erasing handwritten notes in medical records because you'd rather not decipher the doctor's handwriting!

In ***Williams v. Sprint/United Mgmt Co.*****,** 230 F.R.D. 640 (D. Kan. 2005), concerns about privileged metadata prompted the defendant to strip out metadata from the native-format spreadsheet files it produced in discovery. The court responded by ordering production of all metadata as maintained in the ordinary course of business, save only privileged and expressly protected metadata.

The court was right to recognize that privileged information need not be produced, wisely distinguishing between surgical redaction and blanket excision. One is redaction following examination of content and a reasoned judgment that particular matters are privileged. The other excises data in an overbroad and haphazard fashion, grounded only on an often-unwarranted concern that the data pared away ***might*** contain privileged information. The baby goes out with the bathwater. Moreover, blanket redaction based on privilege concerns doesn't relieve a party of the obligation to log and disclose such redaction. The defendant in *Williams* not only failed to examine or log items redacted, it left it to the plaintiff to figure out that something was missing.

The underlying principle is that the requesting party is entitled to the metadata benefits available to the producing party. That is, the producing party may not vandalize or hobble electronic evidence for production without adhering to the same rules attendant to redaction of privileged and confidential information from paper documents.

**The requesting party is entitled to the metadata benefits available to the producing party.**

### Resolve Production Issues

Like other forms of electronic evidence, metadata may be produced in its native and near-native formats, as a database or a delimited load file, in an image format, hosted in an online database or even furnished as a paper printout. However, metadata presents more daunting production challenges than other electronic evidence. One hurdle is that metadata is often unintelligible outside its native environment without processing and labeling. How can you tell if an encoded value describes the date of creation, modification or last access without both decoding the value *and* preserving its significance with labels? Another issue is that metadata isn't always textual. It may consist of no more than a flag in an index entry—just a one or zero—wholly without meaning unless you know what it denotes. A third challenge to producing metadata lies in finding ways to preserve the relationship between metadata and the data it describes and, when obliged to do so, to present both the data and metadata so as to be electronically searchable.

When files are separated from their metadata, we lose much of the ability to sort, manage and authenticate them. Returning to the voice mail example, unless the sound component of the message (e.g., the WAV file) is paired with its metadata, a reviewer must listen to the message in real time, hoping to identify the voice and deduce the date of the call from the message. It's a Herculean task without metadata, but a task made much simpler if the producing party, *e.g.,* drops the WAV file into an Adobe PDF file as an embedded sound file, then inserts the metadata in the image layer. Now, a reviewer can both listen to the message and search and sort by the metadata.

Sometimes, simply producing a table, spreadsheet or load file detailing originating metadata values will suffice. On other occasions, only native production or forensically qualified imaging will suffice to carry forward relevant metadata. Determining the method of metadata production best suited to the case demands planning, guidance from experts and cooperation with the other side.

### Beyond Data about Data

The world's inexorable embrace of digital technologies serves to escalate the evidentiary and functional value of metadata in e-discovery. Today, virtually all information is born electronically, bound to and defined by its metadata as we are bound to and defined by our DNA. The proliferation and growing importance of metadata dictates that we move beyond unhelpful definitions like "data about data," toward a fuller appreciation of metadata's many forms and uses.

**Today, virtually all information is born electronically, bound to and defined by its metadata as we are bound to and defined by our DNA.**

## Appendix A: Just Ones and Zeroes

SECTOR
512 bytes

The binary data above comprises a single hard drive sector storing a binary encoding of the text below (excerpted from *David Copperfield* by Charles Dickens):

> I was born with a caul, which was advertised for sale, in the newspapers, at the low price of fifteen guineas. Whether sea-going people were short of money about that time, or were short of faith and preferred cork jackets, I don't know; all I know is, that there was but one solitary bidding, and that was from an attorney connected with the bill-broking business, who offered two pounds in cash, and the balance in sherry, but declined to be guaranteed from drowning on any higher bargain. Consequently the advertisement was withdrawn at a dead loss - for as to sherry, my poor dear mother's own sherry was in the market then - and ten years afterwards, the caul was put up in a raffle down in our part of the country, to fifty members at half-a-crown a head, the winner to spend five shillings. I was present myself, and I remember to have felt quite uncomfortable and confused, at a part of myself being disposed of in that way. The caul was won, I recollect, by an old lady with a hand-basket, who, very reluctantly, pr

A 1 terabyte hard drive ($75 at your local Wal-Mart) contains more than *2.1 billion* sectors.

## Appendix B: Windows Shell Item Properties

| Category | Properties | |
|---|---|---|
| Audio | ChannelCount<br>Compression<br>EncodingBitrate<br>Format<br>IsVariableBitRate | PeakValue<br>SampleRate<br>SampleSize<br>StreamName<br>StreamNumber |
| Calendar | Duration<br>IsOnline<br>IsRecurring<br>Location<br>OptionalAttendeeAddresses<br>OptionalAttendeeNames<br>OrganizerAddress<br>OrganizerName | ReminderTime<br>RequiredAttendeeAddresses<br>RequiredAttendeeNames<br>Resources<br>ResponseStatus<br>ShowTimeAs<br>ShowTimeAsText |
| Communication | AccountName<br>DateItemExpires<br>FollowupIconIndex<br>HeaderItem<br>PolicyTag | SecurityFlags<br>Suffix<br>TaskStatus<br>TaskStatusText |
| Computer | DecoratedFreeSpace | |
| Contact | Anniversary<br>AssistantName<br>AssistantTelephone<br>Birthday<br>BusinessAddress<br>BusinessAddressCity<br>BusinessAddressCountry<br>BusinessAddressPostalCode<br>BusinessAddressPostOfficeBox<br>BusinessAddressState<br>BusinessAddressStreet<br>BusinessFaxNumber<br>BusinessHomePage<br>BusinessTelephone<br>CallbackTelephone<br>CarTelephone<br>Children<br>CompanyMainTelephone<br>Department<br>EmailAddress<br>EmailAddress2<br>EmailAddress3<br>EmailAddresses<br>EmailName<br>FileAsName<br>FirstName<br>FullName<br>Gender<br>GenderValue<br>Hobbies<br>HomeAddress<br>HomeAddressCity<br>HomeAddressCountry<br>HomeAddressPostalCode<br>HomeAddressPostOfficeBox<br>HomeAddressState | HomeTelephone<br>IMAddress<br>Initials<br>JA.CompanyNamePhonetic<br>JA.FirstNamePhonetic<br>JA.LastNamePhonetic<br>JobTitle<br>Label<br>LastName<br>MailingAddress<br>MiddleName<br>MobileTelephone<br>NickName<br>OfficeLocation<br>OtherAddress<br>OtherAddressCity<br>OtherAddressCountry<br>OtherAddressPostalCode<br>OtherAddressPostOfficeBox<br>OtherAddressState<br>OtherAddressStreet<br>PagerTelephone<br>PersonalTitle<br>PrimaryAddressCity<br>PrimaryAddressCountry<br>PrimaryAddressPostalCode<br>PrimaryAddressPostOfficeBox<br>PrimaryAddressState<br>PrimaryAddressStreet<br>PrimaryEmailAddress<br>PrimaryTelephone<br>Profession<br>SpouseName<br>Suffix<br>TelexNumber<br>TTYTDDTelephone |

| | | |
|---|---|---|
| | HomeAddressStreet<br>HomeFaxNumber | WebPage |
| Devices | Device.PrinterURL<br>DeviceInterface.PrinterDriverDirectory<br>DeviceInterface.PrinterDriverName<br>DeviceInterface.PrinterName<br>DeviceInterface.PrinterPortName<br>BatteryLife<br>BatteryPlusCharging<br>BatteryPlusChargingText<br>Category<br>CategoryGroup<br>CategoryPlural<br>ChargingState<br>Connected<br>ContainerId<br>DefaultTooltip<br>DeviceDescription1<br>DeviceDescription2<br>DiscoveryMethod<br>FriendlyName<br>FunctionPaths<br>InterfacePaths<br>IsDefault<br>IsNetworkConnected<br>IsShared<br>IsSoftwareInstalling<br>LaunchDeviceStageFromExplorer<br>LocalMachine<br>Manufacturer<br>MissedCalls | ModelName<br>ModelNumber<br>NetworkedTooltip<br>NetworkName<br>NetworkType<br>NewPictures<br>Notification<br>Notifications.LowBattery<br>Notifications.MissedCall<br>Notifications.NewMessage<br>Notifications.NewVoicemail<br>Notifications.StorageFull<br>Notifications.StorageFullLinkText<br>NotificationStore<br>NotWorkingProperly<br>Paired<br>PrimaryCategory<br>Roaming<br>SafeRemovalRequired<br>SharedTooltip<br>SignalStrength<br>Status1<br>Status2<br>StorageCapacity<br>StorageFreeSpace<br>StorageFreeSpacePercent<br>TextMessages<br>Voicemail |
| Digital Rights Mgmt. | DatePlayExpires<br>DatePlayStarts<br>Description | IsProtected<br>PlayCount |
| Document | ByteCount<br>CharacterCount<br>ClientID<br>Contributor<br>DateCreated<br>DatePrinted<br>DateSaved<br>Division<br>DocumentID<br>HiddenSlideCount<br>LastAuthor<br>LineCount<br>Manager | MultimediaClipCount<br>NoteCount<br>PageCount<br>ParagraphCount<br>PresentationFormat<br>RevisionNumber<br>Security<br>SlideCount<br>Template<br>TotalEditingTime<br>Version<br>WordCount |
| GPS | Altitude<br>AltitudeDenominator<br>AltitudeNumerator<br>AltitudeRef | ImgDirection<br>ImgDirectionDenominator<br>ImgDirectionNumerator<br>ImgDirectionRef |

| | | |
|---|---|---|
| | AreaInformation | Latitude |
| | Date | LatitudeDenominator |
| | DestBearing | LatitudeNumerator |
| | DestBearingDenominator | LatitudeRef |
| | DestBearingNumerator | Longitude |
| | DestBearingRef | LongitudeDenominator |
| | DestDistance | LongitudeNumerator |
| | DestDistanceDenominator | LongitudeRef |
| | DestDistanceNumerator | MapDatum |
| | DestDistanceRef | MeasureMode |
| | DestLatitude | ProcessingMethod |
| | DestLatitudeDenominator | Satellites |
| | DestLatitudeNumerator | Speed |
| | DestLatitudeRef | SpeedDenominator |
| | DestLongitude | SpeedNumerator |
| | DestLongitudeDenominator | SpeedRef |
| | DestLongitudeNumerator | Status |
| | DestLongitudeRef | Track |
| | Differential | TrackDenominator |
| | DOP | TrackNumerator |
| | DOPDenominator | TrackRef |
| | DOPNumerator | VersionID |
| Image | BitDepth | Dimensions |
| | ColorSpace | HorizontalResolution |
| | CompressedBitsPerPixel | HorizontalSize |
| | CompressedBitsPerPixelDenominator | ImageID |
| | CompressedBitsPerPixelNumerator | ResolutionUnit |
| | Compression | VerticalResolution |
| | CompressionText | VerticalSize |
| Journal | Contacts | EntryType |
| Link | Arguments | Status |
| | Comment | TargetExtension |
| | DateVisited | TargetParsingPath |
| | Description | TargetSFGAOFlags |
| Media | AuthorUrl | FrameCount |
| | AverageLevel | MCDI |
| | ClassPrimaryID | MetadataContentProvider |
| | ClassSecondaryID | Producer |
| | CollectionGroupID | PromotionUrl |
| | CollectionID | ProtectionType |
| | ContentDistributor | ProviderRating |
| | ContentID | ProviderStyle |
| | CreatorApplication | Publisher |
| | CreatorApplicationVersion | SubscriptionContentId |
| | DateEncoded | SubTitle |
| | DateReleased | UniqueFileIdentifier |
| | Duration | UserNoAutoInfo |
| | DVDID | UserWebUrl |
| | EncodedBy | Writer |
| | EncodingSettings | Year |

| | | |
|---|---|---|
| Message | AttachmentContents | FromName |
| | AttachmentNames | HasAttachments |
| | BccAddress | IsFwdOrReply |
| | BccName | MessageClass |
| | CcAddress | ProofInProgress |
| | CcName | SenderAddress |
| | ConversationID | SenderName |
| | ConversationIndex | Store |
| | DateReceived | ToAddress |
| | DateSent | ToDoFlags |
| | Flags | ToDoTitle |
| | FromAddress | ToName |
| Music | AlbumArtist | Genre |
| | AlbumID | InitialKey |
| | AlbumTitle | IsCompilation |
| | Artist | Lyrics |
| | BeatsPerMinute | Mood |
| | Composer | PartOfSet |
| | Conductor | Period |
| | ContentGroupDescription | SynchronizedLyrics |
| | DisplayArtist | TrackNumber |
| Note | Color | ColorText |
| Photo | Aperture | FocalPlaneXResolutionDenominator |
| | ApertureDenominator | FocalPlaneXResolutionNumerator |
| | ApertureNumerator | FocalPlaneYResolution |
| | Brightness | FocalPlaneYResolutionDenominator |
| | BrightnessDenominator | FocalPlaneYResolutionNumerator |
| | BrightnessNumerator | GainControl |
| | CameraManufacturer | GainControlDenominator |
| | CameraModel | GainControlNumerator |
| | CameraSerialNumber | GainControlText |
| | Contrast | ISOSpeed |
| | ContrastText | LensManufacturer |
| | DateTaken | LensModel |
| | DigitalZoom | LightSource |
| | DigitalZoomDenominator | MakerNote |
| | DigitalZoomNumerator | MakerNoteOffset |
| | Event | MaxAperture |
| | EXIFVersion | MaxApertureDenominator |
| | ExposureBias | MaxApertureNumerator |
| | ExposureBiasDenominator | MeteringMode |
| | ExposureBiasNumerator | MeteringModeText |
| | ExposureIndex | Orientation |
| | ExposureIndexDenominator | OrientationText |
| | ExposureIndexNumerator | PeopleNames |
| | ExposureProgram | PhotometricInterpretation |
| | ExposureProgramText | PhotometricInterpretationText |
| | ExposureTime | ProgramMode |
| | ExposureTimeDenominator | ProgramModeText |
| | ExposureTimeNumerator | RelatedSoundFile |

| | | |
|---|---|---|
| | Flash<br>FlashEnergy<br>FlashEnergyDenominator<br>FlashEnergyNumerator<br>FlashManufacturer<br>FlashModel<br>FlashText<br>FNumber<br>FNumberDenominator<br>FNumberNumerator<br>FocalLength<br>FocalLengthDenominator<br>FocalLengthInFilm<br>FocalLengthNumerator<br>FocalPlaneXResolution | Saturation<br>SaturationText<br>Sharpness<br>SharpnessText<br>ShutterSpeed<br>ShutterSpeedDenominator<br>ShutterSpeedNumerator<br>SubjectDistance<br>SubjectDistanceDenominator<br>SubjectDistanceNumerator<br>TagViewAggregate<br>TranscodedForSync<br>WhiteBalance<br>WhiteBalanceText |
| PropGroup | Advanced<br>Audio<br>Calendar<br>Camera<br>Contact<br>Content<br>Description<br>FileSystem<br>General<br>GPS | Image<br>Media<br>MediaAdvanced<br>Message<br>Music<br>Origin<br>PhotoAdvanced<br>RecordedTV<br>Video |
| PropList | ConflictPrompt<br>ContentViewModeForBrowse<br>ContentViewModeForSearch<br>ExtendedTileInfo<br>FileOperationPrompt<br>FullDetails<br>InfoTip | NonPersonal<br>PreviewDetails<br>PreviewTitle<br>QuickTip<br>TileInfo<br>XPDetailsPanel |
| RecordedTV | ChannelNumber<br>Credits<br>DateContentExpires<br>EpisodeName<br>IsATSCContent<br>IsClosedCaptioningAvailable<br>IsDTVContent<br>IsHDContent | IsRepeatBroadcast<br>IsSAP<br>NetworkAffiliation<br>OriginalBroadcastDate<br>ProgramDescription<br>RecordingTime<br>StationCallSign<br>StationName |
| Search | AutoSummary<br>ContainerHash<br>Contents<br>EntryID<br>ExtendedProperties<br>GatherTime<br>HitCount<br>IsClosedDirectory | IsFullyContained<br>QueryFocusedSummary<br>QueryFocusedSummaryWithFallback<br>Rank<br>Store<br>UrlToIndex<br>UrlToIndexWithModificationTime |
| Shell | DescriptionID<br>InfoTipText<br>InternalName | TargetUrl<br>NamespaceCLSID<br>SFGAOFlagsStrings |

| | | |
|---|---|---|
| | TargetSFGAOFlagsStrings | |
| Software | AppUserModel.ExcludeFromShowInNewInstall<br>AppUserModel.ID<br>AppUserModel.IsDestListSeparator<br>AppUserModel.PreventPinning<br>AppUserModel.RelaunchCommand | AppUserModel.RelaunchDisplayNameResource<br>AppUserModel.RelaunchIconResource<br>DateLastUsed<br>ProductName |
| Sync | Comments<br>ConflictDescription<br>ConflictFirstLocation<br>ConflictSecondLocation<br>HandlerCollectionID<br>HandlerID<br>HandlerName | HandlerType<br>HandlerTypeLabel<br>ItemID<br>ItemName<br>ProgressPercentage<br>State<br>Status |
| System | AcquisitionID<br>ApplicationName<br>Author<br>Capacity<br>Category<br>Comment<br>Company<br>ComputerName<br>ContainedItems<br>ContentStatus<br>ContentType<br>Copyright<br>DateAccessed<br>DateAcquired<br>DateArchived<br>DateCompleted<br>DateCreated<br>DateImported<br>DateModified<br>DueDate<br>EndDate<br>FileAllocationSize<br>FileAttributes<br>FileCount<br>FileDescription<br>FileExtension<br>FileFRN<br>FileName<br>FileOwner<br>FileVersion<br>FindData<br>FlagColor<br>FlagColorText<br>FlagStatus<br>FlagStatusText<br>FreeSpace<br>FullText | IsShared<br>ItemAuthors<br>ItemClassType<br>ItemDate<br>ItemFolderNameDisplay<br>ItemFolderPathDisplay<br>ItemFolderPathDisplayNarrow<br>ItemName<br>ItemNameDisplay<br>ItemNamePrefix<br>ItemParticipants<br>ItemPathDisplay<br>ItemPathDisplayNarrow<br>ItemType<br>ItemTypeText<br>ItemUrl<br>Keywords<br>Kind<br>KindText<br>Language<br>LayoutPattern.ContentViewModeForBrowse<br>LayoutPattern.ContentViewModeForSearch<br>MileageInformation<br>MIMEType<br>Null<br>OfflineAvailability<br>OfflineStatus<br>OriginalFileName<br>OwnerSID<br>ParentalRating<br>ParentalRatingReason<br>ParentalRatingsOrganization<br>ParsingBindContext<br>ParsingName<br>ParsingPath |

| | | |
|---|---|---|
| | Identity<br>Identity.Blob<br>Identity.DisplayName<br>Identity.IsMeIdentity<br>Identity.PrimaryEmailAddress<br>Identity.ProviderID<br>Identity.UniqueID<br>Identity.UserName<br>IdentityProvider.Name<br>IdentityProvider.Picture<br>ImageParsingName<br>Importance<br>ImportanceText<br>IsAttachment<br>IsDefaultNonOwnerSaveLocation<br>IsDefaultSaveLocation<br>IsDeleted<br>IsEncrypted<br>IsFlagged<br>IsFlaggedComplete<br>IsIncomplete<br>IsLocationSupported<br>IsPinnedToNameSpaceTree<br>IsRead<br>IsSearchOnlyItem<br>IsSendToTarget | PerceivedType<br>PercentFull<br>Priority<br>PriorityText<br>Project<br>ProviderItemID<br>Rating<br>RatingText<br>Sensitivity<br>SensitivityText<br>SFGAOFlags<br>SharedWith<br>ShareUserRating<br>SharingStatus<br>OmitFromView<br>SimpleRating<br>Size<br>SoftwareUsed<br>SourceItem<br>StartDate<br>Status<br>Subject<br>Thumbnail<br>ThumbnailCacheId<br>ThumbnailStream<br>Title<br>TotalFileSize<br>Trademarks |
| Task | BillingInformation<br>CompletionStatus | Owner |
| Video | Compression<br>Director<br>EncodingBitrate<br>FourCC<br>FrameHeight<br>FrameRate<br>FrameWidth | HorizontalAspectRatio<br>SampleSize<br>StreamName<br>StreamNumber<br>TotalBitrate<br>TranscodedForSync<br>VerticalAspectRatio |
| Volume | FileSystem<br>IsMappedDrive | IsRoot |

# Appendix C: Outlook Standard Properties

% Complete
Account
Actual Work
Address Selected
Address Selector
All Day Event
Anniversary
Assigned
Assistant's Name
Assistant's Phone
Attachment
Bcc
Billing Information
Birthday
Business Address
Business Address City
Business Address Country
Business Address PO Box
Business Address Postal Code
Business Address State
Business Address Street
Business Fax
Business Home Page
Business Phone
Business Phone 2
Callback
Car Phone
Categories
Cc
Changed By
Children
City
Color
Company
Company
Company Main Phone
Complete
Computer Network Name
Contact
Contacts
Content
Conversation
Country
Created
Customer ID
Date Completed
Defer until
Department
Distribution List Name
Do Not AutoArchive
Download State
Due By
Due Date
Duration
E-mail
E-mail 2
E-mail 3
E-mail Selected
E-mail Selector
End
Entry Type
Expires
File As
First Name
Flag Status
Follow-up Flag
From
FTP Site
Full Name
Gender
Government ID Number
Have Replies Sent To
Hobbies
Home Address
Home Address City
Home Address Country
Home Address PO Box
Home Address Postal Code
Home Address State
Home Address Street
Home Fax
Home Phone
Home Phone 2
Icon
Importance
In Folder
Initials
Internet Free Busy Address
ISDN
Job Title
Journal
Junk E-Mail Type
Language
Last Name
Last Saved Time Location
Mailing Address
Mailing Address Indicator
Manager's Name
Meeting Status
Message
Message Class
Message Flag
Middle Name
Mileage
Mobile Phone
Modified
Meeting Workspace URL
Nickname
Notes
Office Location
Optional Attendees
Organizational ID Number
Organizer
Other Address
Other Address City
Other Address Country
Other Address PO Box
Other Address Postal Code
Other Address State
Other Address Street
Other Fax
Other Phone
Outlook Internal Version
Outlook Version
Owner
Pager
Personal Home Page
Phone n Selected
Phone n Selector
PO Box
Primary Phone
Priority
Private
Profession
Radio Phone
Read
Received
Recurrence
Recurrence Pattern
Recurrence Range End
Recurrence Range Start
Recurring
Referred By
Remind Beforehand
Reminder
Reminder Override Default
Reminder Sound
Reminder Sound File
Reminder Time
Reminder Topic
Remote Status
Request Status
Requested By
Required Attendees
Resources
Response Requested
Retrieval Time
Role
Schedule+ Priority
Send Plain Text Only
Sensitivity
Sent
Show Time As
Size
Spouse
Start
Start Date
State
Status
Street Address
Subject
Suffix
Team Task
Telex
Title
To
Total Work
Tracking Status
TTY/TDD Phone
User Field 1
User Field 2
User Field 3
User Field 4
Web Page
ZIP/Postal Code

## Appendix D: Exemplar Metadata Protocol: Native Production

*The following protocol is an example of how one might designate forms and fields for production. It was drafted for use in a particular case and geared to the needs and capabilities of particular parties. It is not a blueprint for other cases, and its language and approach should be emulated only when careful analysis suggests so doing is likely to be effective, economical and proportionate, as well as consistent with applicable law and rules of practice. This Protocol contemplates principally native production and the participation of an ESI Special Master.*

*Note that in addressing metadata and load files in section VI, this Protocol distinguishes between metadata fields applicable to all ESI, to e-mail and attachments and to imaged paper documents. It's important to recognize that there is not one omnibus complement of metadata applicable to all forms of ESI. You must identify and select the fields with particular relevance and utility for your case and applicable to the particular types and forms of ESI produced. See "Metadata Musts" at p. 11, supra.*

*Note also that names assigned to the load file fields are arbitrary. How one names fields in load files is largely immaterial so long as the field name chosen is unique. In practice, when describing the date an e-mail was sent, some label the field "Sent_Date," others use "Datesent" and still others use "Date_Sent." There is no rule on this, nor need there be. What matters is that the information that will be used to populate the field be clearly and unambiguously defined and not be unduly burdensome to extract. Oddly, the e-discovery industry has not settled upon a standard naming convention for these metadata fields.*

### I. General Provisions

**a.** "**Information items**" as used here are individual documents and records (including associated metadata) whether stored on paper or film, as a discrete "file" stored electronically, optically or magnetically or as a record within a database, archive or container file.

**b.** "**Data sources**" as used here are locations or media used to house information items. Data sources include places, like warehouses, file rooms, file cabinets, bankers boxes and folders, as well as storage media like hard drives, thumb drives, diskettes, optical disks, handheld devices, online storage, file shares (e.g., allocated network storage), databases (including e-mail client applications and servers) and container files (e.g., PST, OST and NSF files, Zip archives and PDF portfolios).

**c.** A party is obliged to consider for preservation and identification all *potentially responsive* information items and data sources over which the party (including its employees, officers and directors) has possession, physical custody or a right or ability to exert direction or control. The belief that an employee, agent or contractor may fail to act upon or conform to an exercise of direction or control is not a justification for any party to fail to undertake an exercise of direction or control directed to preservation, identification and, as potentially responsive or relevant, production of information items and data sources.

**d.** To reduce cost and preserve maximum utility of electronically stored information, **production in native electronic formats shall be the required form of ESI production in response to discovery in this cause**. Unless this ESI Production Protocol expressly permits or requires conversion, or the Court or the ESI Special Master expressly permit same on application from a party, native electronic file formats shall not be converted from their usual and customary format to other formats for production.

**e.** Production of information in electronic formats shall not relieve the Producing Party of the obligation to act with reasonable diligence to preserve the native electronic data sources of the information items produced and relevant metadata. Parties should be vigilant not to wipe or dispose of source media while under a preservation duty.

**f.** Where, in the usual course of business a paper document has been imaged as a TIFF or PDF file or an electronic file has been printed or converted to an imaged format, the electronic file and the hard copy or imaged counterpart *may not be treated as identical or cumulative* such that the paper/imaged format may be produced in lieu of its electronic source.

**g.** ***A party must produce OCR textual content when the party produces images in lieu of paper originals.*** A party is not obligated to undertake optical character recognition (OCR) for information items in the PDF, TIFF or JPG file formats when the party does not hold the textual contents of such files in an electronically searchable form at the time of the request for production; provided, however, ***that if a party adds or acquires such electronic searchability, the party adding or acquiring same shall supplement production to make such capability available to other parties in a reasonable manner at the producing party's cost.***

## II. Scope

**a.** Parties must act with reasonable diligence to identify and produce responsive, non-privileged ESI under their care, in their custody or subject to their control, notwithstanding its location, format or medium. Any party may, upon application for relief, seek to limit this duty by showing that compliance would impose upon the party an undue burden or cost that is disproportionate to the claims asserted by or against the party.

**b.** Parties shall produce **responsive, non-privileged** files with the following extensions: ace, arc, arj, arx, bh, cal, cat, csv, dat, db, dbx, doc, docx, dot, dotm, dotx, dst, dstx, efx, email, eml, gnu, gra, gz, gzip, jar, lha, mbox, mbx, mdb, mde, mde, mpp, msg, nsf, ost, pdf, pnf, pot, potx, ppt, pptx, pst, pub, rar, rpt, rtf, shw, tar, taz, thmx, tif, tiff, txt, tz, wbk, wk4, wmf, wpd, wps, wri, xla, xlam, xlm, xls, xlsm, xlsx, xlt, xltx, z, zip.

**c.** Parties shall further **produce responsive, non-privileged** files with the following extensions : ac$, atc, avi, bmp, cdd, cdr, dsp, dtd, dwf, dwg, dws, dwt, dxf, gif, hdi, jpeg, jpg, mov, mp3, mp4, mpeg, mpg, p3, pdd, plt, png, prx, rfa, shx, vsd, vss, vst, wav, wpg.

*Note: Though frequently large in size, these file types rarely contain significant textual content and so rarely warrant use of the more costly methods needed to process text-rich file formats for production.*

**d.** The file types listed in a. and b. above must be identified and considered in searching for potentially responsive ESI; however, this is not an exclusive list of all potentially responsive file types. Parties are expected to apply their knowledge of information systems and applications used by employees when seeking to identify potentially responsive file types, as well as to make a reasonably diligent search for responsive file types not listed.

**e.** Parties may exclude from production all files customarily found in folders reserved to “Temporary Internet Files” with the exception of the contents of Outlook “OLK” folders.

## III. Unique Production Identifier (UPI)

**a.** Other than paper originals and images of paper documents (including redacted ESI), no ESI produced in discovery need be converted to a paginated format nor embossed with a Bates number.

**b.** Each item of ESI (e.g., native file, document image or e-mail message) shall be identified by naming the item to correspond to a Unique Production Identifier according to the following protocol:

i. The first four (4) characters of the filename will reflect a unique alphanumeric designation identifying the party making production;

ii. The next nine (9) characters will be a unique, sequential numeric value assigned to the item by the producing party. This value shall be padded with leading zeroes as needed to preserve its length;

iii. The final five (5) characters are reserved to a sequence beginning with a dash (-) followed by a four digit number reflecting pagination of the item when printed to paper or converted to an image format for use in proceedings or when attached as exhibits to pleadings.

iv. By way of example, a Microsoft Word document produced by Acme in its native format might be named: ACME000000123.doc. Were the document printed out for use in deposition, page six of the printed item must be embossed with the unique identifier ACME000000123-0006.

## IV. E-Mail and Attachments

### a. Form of Production

**i.** E-mail from Exchange/Outlook environments will be produced in the single message MAPI-compliant MSG format. Messages in webmail, Lotus Notes or GroupWise formats shall be converted to the MSG format with care taken to, as feasible, preserve as discrete fields the information listed in VI(c)(x) through (xxi).

**ii.** Attachments to e-mails need not be produced separately from the transmitting message when the attachments are embedded within the message as MIME-compliant, Base-64 content (or other encoding scheme employed in the native source).

## V. Databases

**a.** Prior to production of any database, the parties will meet and confer regarding the reasonableness and feasibility of any request for production of or from a database including the scope, form and content of any such production. Producing Parties shall make reasonable efforts to produce responsive information and data from databases using existing query and reporting capabilities of the database software. To facilitate such interrogation, the Producing Party shall act reasonably and cooperatively to comply with requests from the Requesting Party for information about the reporting capabilities, structure, organization, query language and schema of the database**.**

**b.** If, after good faith efforts to reach agreement, the parties cannot agree, either party may seek assistance from the ESI Special Master regarding the discoverability, form, and scope of production of data from a database.

## VI. Load Files

**a.** Consistent with the provisions of this Protocol, each party shall produce responsive ESI to other parties in its native electronic format accompanied by a load file in one of the following formats to be designated in writing by the requesting party:

**i.** Concordance;
**ii.** Summation;
**iii.** Ringtail; or
**iv.** Such other format as the Requesting and Producing Parties may agree upon.

**b.** **Designation by Party:** Each party shall designate their preferred load file format by no later than 15 days from the entry of an order adopting this ESI protocol. If a party fails to make a timely designation, the party will be deemed to have selected the Concordance load file format.

**c.** **Required Fields:** Any load file format employed shall include the following delimited fields:

**i.** **Identifier** – UPI- The unique production identifier of the item;
**ii.** **Source Name** – NAME - The original name of the item or file when collected from the source custodian or system;
**iii.** **MD5 Hash** – MD5 - The MD5 hash value of the item as produced;
**iv.** **Custodian** – CUSTODIAN -The unique identifier for the original custodian or source system from which the item was collected;

v. **Source Path** – SPATH -The fully qualified file path from root of the location from which the item was collected;

vi. **Production Path** – NATIVELINK -The file path to the item from the root of the production media;

vii. **Modified Date** – MODDATE -The last modified date of the item when collected from the source custodian or system;

viii. **Modified Time** – MODTIME - The last modified time of the item when collected from the source custodian or system; and

ix. **UTC Offset** – UTCOFF - The UTC/GMT offset of the item's modified date and time.

**The following additional fields shall accompany production of e-mail messages:**

x. **To** – RECIPIENT – Addressee(s) of the message;

xi. **From** – FROM – The e-mail address of the person sending the message;

xii. **CC** – CC – Person(s) copied on the message;

xiii. **BCC** – BCC – Person(s)blind copied on the message;

xiv. **Date Sent** – DATESENT – date the message was sent;

xv. **Time Sent** – TIMESENT - time the message was sent;

xvi. **Subject** – SUBJECT – Subject line of the message;

xvii. **Date Received** – DATERCVD – date the message was received;

xviii. **Time Received** – TIMERCVD - time the message was received;

xix. **Attachments** – ATTACHMENTID - The beginning UPI(s) of attachments, delimited by comma

xx. **Mail Folder Path** – MAILPATH - the path of the message from the root of the mail folder;

xxi. **Message ID** – MESSAGEID – The Microsoft Outlook or similar unique message identifier.

**The following additional fields shall accompany images of paper documents:**

xxii. **Beginning Identifier** – BEGNO – The beginning unique production identifier for the first page of the document;

xxiii. **Ending Identifier** – ENDNO – The ending unique production identifier for the first page of the document;

xxiv. **Page Count** – PGCOUNT -The total number of pages in the document;

xxv. **Location** – LOCATION – The source box or other location identifier needed to trace the document to its source.

d. **Optical Character Recognition Load Files**

i. **E-Mail Messages and Attachments:** Text shall be extracted from native email message bodies and produced as document-level text files (.txt) named to correspond to the UPI of the message and organized into separate folder(s) from the messages. The full text of attachments to messages need not be extracted;

however, if it is extracted, it must be produced as document-level text files (.txt) named to correspond to the UPI of the attachments and organized into separate folder(s) from the messages.

- **ii. Document Images:** Text extracted from paper document images using optical character recognition shall be produced as document-level text files (.txt) named to correspond to the UPI of the image and organized into separate folder(s) from the images.

## VII. Deduplication

**a.** Producing parties should not globally (i.e. horizontally) deduplicate production or apply e-mail threading without express authorization from the ESI Special Master or the Court. Producing parties are encouraged to deduplicate vertically; that is, within a single custodian's document set.

## VIII. Redaction

**a.** Files that must be redacted should be produced as TIFF or PDF images accompanied by load files conforming to Paragraph VI (above) and containing extracted text acquired after redaction (i.e., excluding redacted content).

**b.** Parties may employ native redaction techniques so long as the method of redaction employed does not significantly impair the usability or searchability of the redacted item and the fact of alteration is disclosed.

**c.** Redactions must be logged in the manner of any other responsive material withheld on claims of privilege or confidentiality.

## IX. Imaging and Unitization

a. **Imaging:** Paper documents produced electronically should be imaged at 300 DPI as single-page Group IV TIFF or multipage PDF files;

- **i.** Imaged documents should be produced with OCR text;
- **ii.** The images shall reflect, without visual degradation, the full and complete information contained in the original document; and,
- **iii.** If an original document contains color used in a manner material to the import or understanding of the document, the Producing Party shall comply with reasonable requests for color image formats of the original at the producing Party's cost. Color documents should be produced as PDF or JPG files.

b. **Unitization:** If a scanned document is more than one page, the unitization of the document and any attachments shall be maintained as it existed in the original when creating the image file. For documents that contain affixed notes, the pages will be scanned both with and without the notes and those pages will be treated as part of the same document. The relationship of documents in a document collection (e.g., cover letter and enclosures, email and attachments, binder containing multiple documents, or other documents where a parent-child relationship exists between the documents) shall be maintained through the scanning or conversion process. If more than one level of parent-child relationship exists, documents will be kept in order, but all will be treated as children of the initial parent document.

## X. Cooperation and Transparency

**a.** Parties shall cooperate to identify and facilitate access to the contents of encrypted, corrupted or difficult-to-access files produced. Parties should work cooperatively to fashion reasonable, precise and cost-effective search strategies and to agree upon and implement appropriate measures for quality assurance and quality control. Parties are obliged to be forthcoming and transparent in disclosing their use of mechanized tools to cull responsive data and encouraged to bring technically-adept personnel together to resolve e-discovery issues.

## XI. Production Media and Transmittal

**a.** Production volumes below 12 gigabytes may be produced on three or fewer DVD-ROM optical disks for Windows-compatible personal computers.

**b.** Production volumes greater than 12 gigabytes shall be produced as uncompressed data on Windows-compatible external hard disk drives employing the USB 2.0 interface.

**c.** Production media shall be prominently and uniquely labeled to identity the Producing Party, the date of production and the range of Unique Production Identifiers on the medium.

**d.** Parties shall accompany the production with a written transmittal that sets out the:

**i.** Range of Unique Item Identifiers included within the production;

**ii.** The total number of items produced;

**iii.** The structure of the load file, including the delimiters employed and a list of delimited fields in the order in which they are employed in the load file;

**iv.** Confirmation that the number of items produced matches the load file(s); and

**v.** A privilege log identifying any documents withheld from the data sources being produced, if any.

**e.** Parties should exercise care to guard against truncation or data loss attendant to overlong file names and file paths.

## XII. Modification

**a.** Any agreement between parties to depart from the requirements of this Protocol as between those parties must be memorialized in a writing signed by counsel for all parties to the agreement and promptly furnished to all parties and to the ESI Special Master. Such agreement does not relieve those parties of their obligation to other parties and the Court pursuant to this Protocol.

## XIII. Procedure for Amending or Obtaining Relief from the ESI Production Protocol.

a. **Amendment:** Any party may request that this ESI Production Protocol be amended. All such requests shall be in writing and submitted to the ESI Special Master for consideration, with a copy of the request served to all parties. Any party may oppose a request to amend the ESI Production Protocol by submitting a written opposition to the ESI Special Master, with a copy of the opposition served to all parties within five days of service of the request to amend. The ESI Special Master shall thereafter issue a written decision on the request.

b. **Relief:** Any party may request relief from any obligation set forth in this ESI Production Protocol. All such requests shall be in writing and submitted to the ESI Special Master for consideration, with a copy of the request served to all parties. Any party may oppose any request for relief by submitting a written opposition to the ESI Special Master, with a copy of the opposition served to all parties, within five days of service of the request for relief. The ESI Special Master thereafter shall issue a written decision.

## About the Author

**CRAIG BALL**
**Trial Lawyer & Special Master**
**Computer Forensic Examiner**
**Author and Educator**

**3723 Lost Creek Blvd.**
**Austin, Texas 78735**
**E-mail: craig@ball.net**
**Web: craigball.com**

**Lab: 512-514-0182**
**Mobile: 713-320-6066**

Craig Ball is a Board Certified trial lawyer, certified computer forensic examiner and electronic evidence expert He's dedicated his career to teaching the bench and bar about forensic technology and trial tactics. After decades trying lawsuits, Craig limits his practice to service as a court-appointed special master and consultant in computer forensics and e-discovery. A prolific contributor to educational programs worldwide--having delivered over 650 presentations and papers--Craig's articles on forensic technology and electronic discovery frequently appear in the national media. He writes a monthly column on computer forensics and e-discovery for Law Technology News and Law.com called "Ball in your Court," recognized as the 2007 and 2008 Gold Medal honoree as “Best Regular Column” as awarded by Trade Association Business Publications International, the 2009 Gold and 2007 Silver Medalist honoree of the American Society of Business Publication Editors as “Best Contributed Column” and the 2006 Silver Medalist honoree as “Best Feature Series” and “Best Contributed Column." Craig Ball has served as the Special Master or testifying expert on computer forensics and electronic discovery in some of the most challenging and best-known cases in the U.S. Named as one of the Best Lawyers in America and a Texas Superlawyer, Craig is a recipient of the Presidents' Award, the State Bar of Texas' most esteemed recognition of service to the profession and of the Bar's Lifetime Achievement Award in Law and Technology.

### EDUCATION

Rice University (B.A., 1979, triple major); University of Texas (J.D., with honors, 1982); Oregon State University (Computer Forensics certification, 2003); EnCase Intermediate Reporting and Analysis Course (Guidance Software 2004); WinHex Forensics Certification Course (X-Ways Software Technology 2005); Certified Data Recovery Specialist (Forensic Strategy Services 2009); numerous other classes on computer forensics and electronic discovery.

### SELECTED PROFESSIONAL ACTIVITIES

Law Offices of Craig D. Ball, P.C.; Licensed in Texas since 1982.
Board Certified in Personal Injury Trial Law by the Texas Board of Legal Specialization
Certified Computer Forensic Examiner, Oregon State University and NTI
Certified Computer Examiner (CCE), International Society of Forensic Computer Examiners
Admitted to practice U.S. Court of Appeals, Fifth Circuit; U.S.D.C., Southern, Northern and Western Districts of Texas.
Member, Editorial Advisory Board, Law Technology News and Law.com (American Lawyer Media)
Board Member, Georgetown University Law School Advanced E-Discovery Institute and E-Discovery Academy
Member, Sedona Conference WG1 on Electronic Document Retention and Production
Member, Educational Advisory Board for LegalTech (largest annual legal technology event)
Special Master, Electronic Discovery, numerous federal and state tribunals
Instructor in Computer Forensics and Electronic Discovery, United States Department of Justice
Lecturer/Author on Electronic Discovery for Federal Judicial Center and Texas Office of the Attorney General
Instructor, HTCIA Annual 2010, Cybercrime Summit, 2006, 2007; SANS Instructor 2009
Contributing Editor, EDDUpdate blog
Special Prosecutor, Texas Commission for Lawyer Discipline, 1995-96
Council Member, Computer and Technology Section of the State Bar of Texas, 2003-2010
Chairman: Technology Advisory Committee, State Bar of Texas, 2000-02
President, Houston Trial Lawyers Association (2000-01); President, Houston Trial Lawyers Foundation (2001-02)
Director, Texas Trial Lawyers Association (1995-2003); Chairman, Technology Task Force (1995-97)
Member, High Technology Crime Investigation Association and International Information Systems Forensics Assn.
Member, Texas State Bar College
Member, Continuing Legal Education Comm., 2000-04, Civil Pattern Jury Charge Comm., 1983-94, State Bar of Texas
Life Fellow, Texas and Houston Bar Foundations
Adjunct Professor, South Texas College of Law, 1983-88