White Paper

Chain of Custody

Authenticating Digital Evidence – Identify and Avoid the Weak Links in Your Chain of Custody

MERRILL LEGAL SOLUTIONS

Contents

Introduction

Chain of custody is a familiar concept in criminal law, but until recent years it was foreign to civil litigators. In the criminal law arena, police would seize evidence, seal it in a plastic bag, label it and sign it into a locked evidence room. If the evidence was taken out for any purpose (for example, for laboratory examination or testing) the withdrawal and its return would be noted on the custody log. Any subsequent removal of the evidence from the locked room would be unlikely until it was presented as evidence at trial.

Historically, evidentiary chain of custody was rarely an issue in civil litigation. The advent of the digital age has made it a major issue because the actual nature of evidence in civil litigation has undergone a radical transformation from tangible paper to electronic data.

Electronic discovery (e-discovery) is a multi-stage process and custody is an issue at every one of those stages.

In the electronic discovery publication, *Arkfeld on Electronic Discovery and Evidence*[1], the author notes the following regarding the importance of chain of custody:

> The purpose of testimony concerning chain of custody is to prove that evidence has not been altered or changed from the time it was collected through production in court. Gallego v. United States of America, 276 F.2d 914 (9th Cir. 1960) (citing United States v. S.B. Panicky & Co., 136 F.2d 413, 415 (2d Cir. 1943)). Chain of custody testimony would include documentation on how the data was gathered, transported, analyzed and preserved for production. This information is important to assist in the authentication of electronic data since it can be easily altered if proper precautions are not taken.

It is also much more complicated to handle electronic data as evidence than it is to sign in tangible narcotics confiscated at the time of arrest and sign them out again at the time of trial. That is because electronic discovery (e-discovery) is a multi-stage process and custody is an issue at every one of those stages.

Stages of electronic discovery[2]

The generally accepted stages of e-discovery are: identification, preservation, collection, processing, review, analysis, production and finally, presentation as evidence during trial.

The requirement to maintain a defensible chain of custody applies to every stage of the e-discovery process. This will involve documenting 1) the methodology used in the forensic acquisition of ESI contained on storage media (such as a hard drive) and 2) the chain of custody of the ESI during and after the retrieval process. The foundation for admissibility of ESI may be attacked by objecting to either aspect of the process.

Because the discovery process has so many steps, the chain of custody can be very long and convoluted. However, the admissibility of ESI will hinge, in part, on laying a proper foundation for the electronic evidence.

[1] Michael R. Arkfeld, Arkfeld on Electronic Discovery and Evidence, § 8.10(C), *Chain of Custody*.

[2] Electronic Discovery or e-discovery is also referred to as discovery of Electronically Stored Information [ESI].

Forensic or non-forensic acquisition

The decision regarding whether ESI collection needs to adhere to strict forensic acquisition methodology should be addressed at the earliest opportunity. If there are concerns regarding preservation of information coupled with concerns regarding data integrity, it may be essential to immediately begin the process to obtain ESI using a forensic methodology. However, if the nature of the documents and their exact contents are not really in dispute (for example, in a contractual dispute where no one contests the terms of the contract) a forensics methodology may not be required. In these situations, a much less rigorous method of document collection may well suffice.

A forensic investigator might photograph a laptop, record its make and model number, its serial number, the date it was seized and the "hash" value or "electronic fingerprint" of the entire hard drive.

Establishing a proper chain of custody for forensic acquisition

In highly contentious litigation, one or both parties may hire computer forensic specialists to locate and seize information for purposes such as identifying assets and establishing proof of wrongdoing. For example, in a contested divorce proceeding, a court may order that the wife be allowed to inspect the husband's personal laptop, especially if there is suspicion concerning possible destruction of electronic evidence. Forensic investigators, acting for the wife's attorney, arrive at the door of the husband's residence with the court order, seize the physical evidence (for instance, the computer laptop), attach copying equipment to it and make a bit-by-bit image of the entire hard drive.

They are not just copying files, but making a forensically exact image that includes everything on the hard drive such as slack space and deleted files that have not yet been overwritten (for example, deleted e-mails to and from a paramour, deleted brokerage accounts, banking records, Internet surfing history and more). Forensics specialists use tools such as Guidance Software's Encase®[3] or Forensic Tool Kit[4]. These tools are designed so that the act of copying the data – an activity that would normally change some information or metadata on the source hard drive – alters nothing on the source hard drive while yielding a mirror image of the information stored.

In this example, the chain of custody documentation required to this point is fairly straightforward:

1. The forensics specialist first takes a photograph of the laptop, known as the original media, and records its make and model number, serial number, the date seized, and very importantly, the "hash" value or "electronic fingerprint" of the entire hard drive.
2. The specialist makes a forensic image of the original media's electronically stored information onto a pristine hard drive. We will call this resulting image the evidentiary copy, sometimes also called the forensic copy.
3. Once copying is complete, a hash value is taken of the evidentiary copy. If the copying process was performed properly, the hash value of the evidentiary copy should match the hash value of the original media – the laptop's hard drive.
4. The original laptop is then returned to the husband.

The original electronic content of the laptop, now stored as the evidentiary copy, is the evidence of interest. The evidentiary copy must be properly preserved to establish

[3] See http://www.guidancesoftware.com/
[4] See: http://www.accessdata.com/common/pagedetail.aspx?PageCode=ftk2test

authenticity. All transfers of custody of the evidentiary copy must be logged. Also at this point, best practices demand that a working copy be made of the evidentiary/forensic copy and that the evidentiary copy be locked away or safeguarded in a secure manner to avoid alteration, damage or spoliation. All subsequent chain of custody logging will now document access to or changes to the working copy – when it was taken to the lab, what files were reviewed, etc.

In this scenario, the tangible and the intangible – the hard drive and the data on it – are almost one and the same. In these situations the courts will accept the ESI stored on the evidentiary copy (with an identical hash value) as virtually indistinguishable from the ESI stored on the original media source drive. The court is likely therefore to accept the foundation for the evidentiary copy and approve its admissibility at trial. This type of forensic acquisition has been effectively employed for years and has resulted in a forensically defensible chain of custody.

The "chain of custody" rule requires that admitted exhibits "be preceded by evidence sufficient to support a finding that the matter in question is what its proponent claims."

Establishing a proper chain of custody for non-forensic acquisition

Many articles written about chain of custody in civil e-discovery presume that initial data extraction is always done in the afore-described forensically exact manner. In fact, this is not the case. Forensic ESI collection is primarily used when there are early issues regarding suspicion of fraud, theft of intellectual property, concealment of assets or other indicia of concealment including deliberate spoliation of ESI.

In the majority of civil litigation such as business contract disputes or tort cases, forensic collection of information is not anticipated. Instead, civil litigation requires preservation, review and production of information relevant to the matters at issue in the litigation, as established in Fed. R. Civ. P. 26(b)(1) which reads:

> Parties may obtain discovery regarding any matter, not privileged, that is relevant to the claim or defense of any party, including the existence, description, nature, custody, condition and location of any books, documents [defined elsewhere in the rules to include "electronically stored information" (ESI)], or other tangible things.

This is the basic scope of the civil discovery rule; nothing here about kicking down any doors or seizing anything.

However, identification and preservation of relevant electronic evidence is only part of the issue. It nonetheless remains extremely important during the collection of this evidence that the chain of custody be kept intact to ensure its admissibility in court. As one author noted[5]:

> Depending on the circumstances of the case, a chain of custody foundation will assist in the admission of evidence. When there is a chance of confusion, or that data may have been altered or tampered with, evidence establishing a chain of custody is important. …The "chain of custody" rule is a variation of the requirement under FED. R. EVID. 901(a) that evidence must be properly authenticated or identified prior to being admitted. *United States v. Turpin*, 65 F.3d 1207, 1213 (4th Cir. 1995) (citations omitted). The "chain of custody" rule requires that admitted exhibits "be preceded by evidence sufficient to support a finding that the matter in question is what its proponent claims."

[5] Michael R. Arkfeld, *Arkfeld on Electronic Discovery and Evidence*, § 5.5 (B), at page 5-40.

In the majority of cases, nobody is interested in what may have been deleted from a hard drive or what a user's Internet surfing history might be. Most of the time, the only thing that is even potentially relevant in litigation is the *active data* on the computers and network file servers and even more likely in the e-mail boxes of the people and departments related to the subject matter of the litigation.

In the normal litigation context, we are not usually talking about busting into a defendant location and seizing computers to make forensically exact copies of hard drives. Nor are we talking about making forensically exact copies of the hard drives of our own personnel's computers. In most civil cases with discoverable electronic data, the data is collected in a more "informal" manner from the custodians' machines, e-mailboxes on the server and the Outlook PST files on the custodian's machine. The data collector may need to move from machine to machine and from server to server, but instead of making a complete forensic copy of every hard drive, the collector makes copies of just those files and folders deemed relevant to the litigation.

Most of the time, the only thing that is even potentially relevant in litigation is the *active data* on the computers and network file servers and even more likely in the e-mail boxes of the people and departments related to the subject matter of the litigation.

Another approach is to use a collection software or appliance to retrieve ESI in a forensic manner without all the steps essential in a traditional forensic collection scenario. These software applications, or appliances, which claim to be able to retrieve ESI in a forensically sound manner without all the steps essential in a traditional forensic collection, begin the collection process by plugging into any location on a computer network. From that location, these tools create a map of all the data on the network and then collect data from any server, desktop, laptop or other device on the network without requiring the data collector to move from one location to another to accomplish the collection. In either case, the person collecting information may in fact aggregate the copied files onto a portable hard drive via a USB connector.

The use of this technology, while seemingly a convenient and less expensive option, is fraught with complications. Most troublesome is when this methodology is attempted by in-house technical staff or by an amateur unversed in the legal implications of data collection.

While this process is "less rigorous" in comparison to the exacting protocols and specialized equipment used in making forensically identical copies of hard drives, it must still comply with all of the basic standards of forensic collection. An electronic fingerprint must be taken to verify that any copies of files are identical to the original files on the user's original media or hard drive. This fingerprint, once again, is our "hash value" as earlier described, except that in this case it is being applied down to the individual file level.

It is critical that the hash value of original files or folders be recorded at the time of acquisition and then matched against the hash value of the copied versions to ensure accurate duplication and maintain the standards essential for authentication. The importance of this step cannot be over-emphasized. The hash value is a critical element throughout the entire chain of custody, in every step of the discovery process. If the hash values on the copies do not match the original, this informal evidentiary copy (proffered to the court for introduction into evidence) may lack sufficient foundation to be admitted.

If at some point there is an objection to the chain of custody record, the hash value of the file being offered as evidence in court – if it is identical to the hash value of the file originally collected from the hard drive of "person X" – will provide one of the necessary foundational elements for admissibility. In addition to the hash value of a file, it is essential to maintain and

Hash Value

What is a hash value? A hash value is an electronic fingerprint. *"A hash value can be applied to a file, a section of a disk, or a whole disk, and recorded. The hash value will change if the data in a file, section or disk is changed or altered."*[6]

One type commonly heard of is the MD-5 Hash Value. Here is an example of what it looks like, generated from a commonplace phrase:
("The quick brown fox jumps over the lazy dog")
= 9e107d9d372bb6826bd81d3542a419d6

Even a small change in the message will result in a completely different hash. For example, changing d to e:
("The quick brown fox jumps over the lazy **e**og")
= ffd93f16876049265fbaef4da268dd0e

Another hash value algorithm is called SHA-1. Instead of 32 characters, it has 40. Here is what it looks like on the same phrase:
("The quick brown fox jumps over the lazy dog")
= 2fd4e1c6 7a2d28fc ed849ee1 bb76e739 1b93eb12

As with the MD-5, even a small change in the message will result in a completely different hash. For example, changing dog to cog:
("The quick brown fox jumps over the lazy **c**og")
= de9f2c7f d25e1b3a fad3e85a 0bd17d9b 100db4b3[7]

The hash value change resulting from the alteration of merely one single character in this sentence results in a change as dramatic as if "War and Peace" had been edited to become "The Cat in the Hat."

The hash value is the critical element that carries through the entire chain of custody, through every step of the discovery process.

provide, upon request, a legally defensible chain of custody log explicitly documenting where, when and how the ESI was preserved, collected and processed.

Ultimately, the admissibility of the ESI will rely upon the sophistication of the collection methodology and the personnel involved in the collection process. If the opposing party objects to the ESI, a complete foundation for the ESI will have to be laid. If the collection process was done in-house, there may be no competent witness available to testify regarding the foundation of the evidence who is not also a party to the litigation. It becomes problematic if the actual client, one of the client's employees or the client's counsel collects the data since they may later be required to establish their credibility and lack of bias in the collection process. If the ESI was collected by someone lacking qualifications or experience, or if the collection methodology does not satisfy the court, the ESI may be excluded from evidence.

[6] These examples of hash values were obtained from Wikipedia: http://en.wikipedia.org/wiki/MD5 and http://en.wikipedia.org/wiki/SHA_hash_functions.
[7] Michael R. Arkfeld, Arkfeld on Electronic Discovery and Evidence, § 8.10(C), *Chain of Custody*

This is a looming trap for any party attempting to introduce electronically stored information. It is easy to see why the process is problematic and should not generally be attempted by in-house technical staff or by an amateur. Although anyone can collect electronic evidence and "hash" the corresponding computer files, it may be prudent to have a third party perform the collection process and generate the hash algorithms.

These issues also make a strong argument for using an electronic evidence service provider or e-discovery vendor who is also qualified in forensic collection methodology. This guarantees that ESI collection will be completed in a forensically defensible manner, as well as provide assurance that if the need for a forensic witness becomes essential, you are ready.

As part of the collection documentation, a collection "manifest" should be created that includes a list of the files collected, their location and their hash values.

Chain of custody log

Once the evidence is properly collected, either using forensic methods or non-forensic methods, it is important to make sure that the chain of custody for that evidence is properly maintained. The foundation for this process is the chain of custody log. Maintaining a proper chain of custody log is essential to a defensible collection method and will assist in laying the foundation for the admission of ESI at trial. The chain of custody log for electronic evidence in civil litigation, in the vast majority of the cases where forensic hard drive imaging is not utilized, will at a minimum contain the following elements.

Information from the collection stage. Whether the process is forensic or non-forensic, you have a requirement to properly maintain your chain of custody.

This includes:

- A description of all devices from which data was copied (including model numbers, serial numbers and locations of each device);
- A description of the folders that were copied; and
- The process used for making the copies.

The collection procedure should create a hash value for each file during the process of reading the file, either by direct copy from the individual machine or over the network. These hash values should be stored, along with a copy of the file in its native format, on the same media onto which the data is copied.

As part of the collection documentation, a collection manifest should be created that includes a list of files collected, their location and their hash values. The manifest can be used to compare the hash values of the collected files and the files produced at the end of the review.

Shipping to the electronic discovery processing service provider. Many data collection services prefer to personally deliver the media holding the collected data directly to the processing company, although commercial realities may dictate that a courier be used. While less than optimal, the tracking capabilities of FedEx and UPS both provide the documentation that comports with what is required to maintain proper chain of custody.

One best practice is to retain an electronic evidence service provider who can perform all of these functions to ensure the integrity of the electronic evidence and avoid "hand-offs", which may add elements of unreliability to the proper maintenance of custody.

Receipt by the electronic discovery processing service provider. The chain of custody log used by the e-discovery processing company should, according to the Electronic Discovery Reference Model (EDRM), contain the following:

- Electronic discovery identification and inventory number (a barcode labeling system is recommended)
- Date received
- Matter name
- Client name
- Client/matter number
- Name of person/company/shipper delivering the evidence
- Description of item(s) (including manufacturer name, model number and unique identifier/serial number whenever possible)
- MD5 Hash of each piece of media where possible (electronic fingerprint)
- Name of person receiving evidence (Logged by......)
- Check Out (check box—Yes/No) If "Yes, Date and Reason
- Custodian name
- Name of recipient (used when evidence is shipped)
- Name of shipper
- Shipper's tracking number
- Date of shipment
- Date of receipt
- Check-in date

At no time should media containing the original evidentiary or working copies of client data be left out in an open, unsecured area or unaccompanied by the person authorized to possess and work with them.

Whenever the originally obtained evidence is accessed, it should only be made available to the small team in charge of logging and securing the evidence. It is clearly preferable to use the working copy of the evidence rather than the evidentiary copy to avoid any damage or spoliation. Any activities involving the evidentiary copy of evidence must be logged as well. After the logging, the owner of the evidence should be sent confirmation that evidence was received.[8]

Processing. The physical premises of the e-discovery processing company should be designed in such a way that media containing collected data are safeguarded in a secure location to which only a select group of people have access. For example, if working copies of evidence are removed from the secured room to load the data onto the processing servers, the media should be signed out, copied to the processing servers as quickly as possible and returned to the locked room without delay. At no time should media containing the original evidentiary or working copies of client data be left out in an open, unsecured area or unaccompanied by the person authorized to possess and work with them. This applies even if the entire company's premises are secured and access is restricted.

Review of evidence. After "native file" processing is complete, the data should once again be checked against the original copies' hash values to verify that the hash values are still identical. If the processing involves changing the native files using TIFF (Tagged Image File Format) images, the data is obviously changed and will then yield an entirely different hash value.

[8] http://edrm.net/wiki/index.php/Processing_-_Audit_and_Chain_of_Custody

At this point, the processed data will either be sent elsewhere – to the law firm for loading on its own internal review system, or to a third party Web host – or it will be loaded onto the e-discovery processing vendor's own Web-hosting application. If the processed data leaves the e-discovery processing vendor's custody and control, again, proper chain-of-custody logs, related both to the media onto which the data was copied and the form of delivery, should be maintained.

Sometimes these deliveries are in the form of electronic transmissions such as files attached to e-mails or files sent via FTP (file transfer protocol) rather than transfer of actual physical media. With any transfer of the data, the chain of custody baton passes to the recipient, whether a law firm or third party Web host. If the processed data stays with the e-discovery vendor for hosting, that vendor's internal chain of custody logs should reflect the handoff from the processing department to the Web hosting department.

When there is a risk of confusion or a risk that ESI may have been altered or tampered with, evidence establishing a chain of custody is crucial to its ultimate use at trial.

Confirmation and production of chain integrity. After the review is complete, and the reviewing party has chosen what electronic documents to produce to the opposing side, the hash algorithms need to be regenerated to verify that what is being produced is identical to that which was first collected. Conversely, this is the stage where the client or counsel will also be receiving production of ESI inbound from the other side. It is essential that you immediately regenerate hash-value algorithms on the incoming data before touching it in any other way. This is a crucial anticipatory action in preparation for the next stage.

Presentation at trial. If you find a smoking gun in your opponent's productions and want to present it at trial, mediation or arbitration, you can prove it is exactly the same as the file you received by comparing its present hash value to the hash fingerprint of it that you made the moment it arrived.

Conversely, if one of your own documents is critical to the proof of your case, you may need to refer back to the hash fingerprint that was first generated during early collection to prove during presentation at trial that it is still the same document, and that nothing has been altered.

The value of outsourcing chain of custody management

Chain of custody is now as important to civil litigators as it is to criminal attorneys. Depending on the circumstances of the case, a chain of custody foundation may have to be rigorously established for the admission of ESI evidence. When there is a risk of confusion or a risk that ESI may have been altered or tampered with, evidence establishing a chain of custody is crucial to its ultimate use at trial.

The first step necessary to establish a consistent chain at trial is generally to ask the custodian or a witness about the origin, storage and handling of the electronic evidence. When faced with a case involving a computer record and ESI evidence, the procedures used to gather the data and how it was copied and preserved will also be placed at issue. In addition, testimony about the storage media used to store and transfer the data, as well as testimony about how the ESI was processed and searched, will also be required.

Laying a proper chain of custody foundation will establish under FED. R. EVID. 901(a) that evidence was properly authenticated or identified prior to being admitted. With a properly established "chain of custody" you can be sure that the evidence will be given its proper weight and consideration by the jury.

Now that civil litigants must be equally cognizant of the necessity of maintaining a proper "chain of custody" to ensure the admissibility of ESI, one must make a critical decision regarding who will be charged with responsibility for the data collection and preservation as well as what methodology will be used. Clearly, these decisions should be made in an informed fashion and should contemplate the potential need for an electronic evidence service provider.

The necessity of ensuring the chain of custody for civil cases has risen in importance and must be carefully planned during the acquisition of your electronic evidence. You may want to consider using a single third party service provider to collect and process the evidence to ensure that standardized procedures are followed. If the collection procedure is challenged, a witness from the electronic evidence service provider can usually offer the relevant testimony setting forth the chain of custody for the electronic evidence and ensure that all the links in the chain of custody are intact.

About Merrill Corporation

Founded in 1968 and headquartered in St. Paul, Minnesota, Merrill Corporation (www.merrillcorp.com) is a leading provider of outsourcing solutions for complex business communication and information management. Merrill's services include document and data management, litigation support, branded communication programs, fulfillment, imaging and printing. Merrill's target markets include the legal, financial services, insurance and real estate industries. With more than 6,300 people in over 70 domestic and 15 international locations, Merrill empowers the communications of the world's leading companies.

About Merrill Legal Solutions

Clients worldwide rely on Merrill Legal Solutions to provide integrated solutions to accurately, cost-effectively, reliably and consistently manage the complex litigation life cycle. Merrill Legal Solutions combines e-discovery expertise and proven solutions with flawless project management and customer service to impact the outcome of litigation, from small to large high-stakes cases. With a single, end-to-end global provider for your litigation support, discovery, deposition services and trial consulting, you will improve your case management while saving time and money.

Corporate Headquarters
One Merrill Circle
St. Paul, MN 55108
800.688.4400
legalsolutions@merrillcorp.com

www.merrillcorp.com/law