doeLEGAL™

Avoiding the Minefields of Litigation Self-Collection

Amended Rule 902 provides a guide to safely traversing the dangers

By Adam Bowers, JD.

WHITEPAPER

© 2018 doeLEGAL, Inc. All rights reserved.

There has been a seismic shift in the Federal Rules of Evidence (FRE); with some far reaching implications. A recent amendment to Rule 902 is causing some corporate entities to question whether they should collect their own litigation data or if they are required to hire a 3rd-party professional. The good news for those corporations is that the amended rule does not necessitate the use of forensic collection experts, but there are still some very important reasons to consider doing just that.

Introduction

All attorneys must minimally understand, and optimally participate in, the identification, preservation, and collection of their clients' ESI (Electronically Stored Information). This white paper provides an overview of self-collection and a cost-benefit analysis of the DIY (do it yourself) model from a lawyer's perspective, not a marketing perspective. This is an important distinction because doeLEGAL is not a forensic collection company, we offer best-of-breed data collection technologies and partner with clients to ensure they know how to be successful. The intent is to take an honest look at the issues surrounding how corporate clients gather and identify litigation data. This white paper is not to be considered legal advice. If you need legal advice, consult with a licensed attorney. With that said, let's look at the changes FRE 902 brought to life and it's impact on collections.

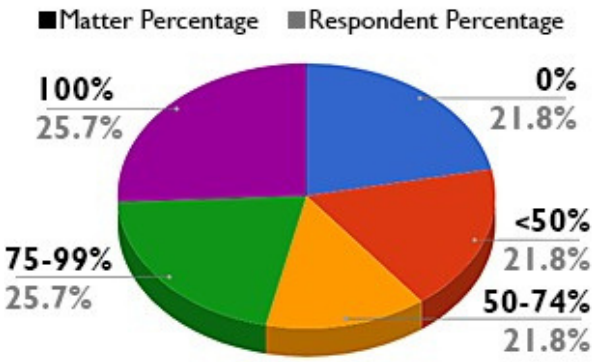
Dangers of allowing employees to conduct legal holds and manage their own data

With all the new technology now available, it is easy to forget that humans are still part of the litigation equation. One case, in particular, exemplifies just how expensive it can be to allow employees to preserve their own litigation data: GN Netcom, Inc. v. Plantronics, Inc. The case has become so well known in eDiscovery circles, much like Zubulake, that it joins the infamous "single name club" - we call it Plantronics. In this case, a senior member of management instructed his sales team to delete certain emails and other data in direct contradiction to a company-instituted legal hold on those email accounts. The intent of the sales manager was seemingly malicious and the Delaware federal judge was not amused.

He sanctioned Plantronics, to the tune of three million USD, plus associated fees and costs. It is important to note that Plantronics was ultimately victorious on the merits of the case, but was still liable for millions in spoliation associated costs: they lost the battle, but won the war. While the Plantronics case serves as an example of intentional destruction of litigation data, there are also less nefarious reasons that employees should not be left to manage their own data preservations or collections.

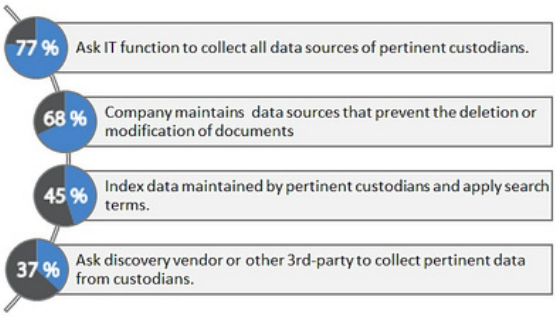
Before we get too far into the concepts of preservation and collection, let's define "custodian self-collection" using the two methods used most. One is where the custodians are responsible for the preservation, searching, and identification of their data while IT would do the actual collection (gathering of ESI). The second is where the custodians are responsible for preserving, searching, identifying, and collecting their own data. Each leads to many of the same challenges and risks, but the first method is far more widely used in business so we will use that as our reference.

Norton Rose Fulbright's 2017 Litigation Trends Annual Survey covered the subject in their section on discovery where they asked respondents: In the past 12 months, for what percentage of your matters have you primarily relied upon self-preservation?



© 2018 doeLEGAL, Inc. All rights reserved.

In a recent webcast1, 62% of the attendees were concerned with spoliation risks associated with custodian self-preservation. However, a survey of corporate counsel, 47% of organizations relied on
custodian self-preservation more than 75% of the time. This shows a clear split between what they know is right and doing what's right. Another question on the Norton Rose Fulbright survey uncovered more about how respondents handled their data preservation: If you do not rely on self-preservation, how do you preserve potentially relevant documents? 77% of respondents who weren't relying on self-preservation were, at least some of the time, employing a "collect everything"
approach to preservation, an approach which creates its own obvious issues such as increased costs, larger volumes of data to review, and potential risk in other legal matters.



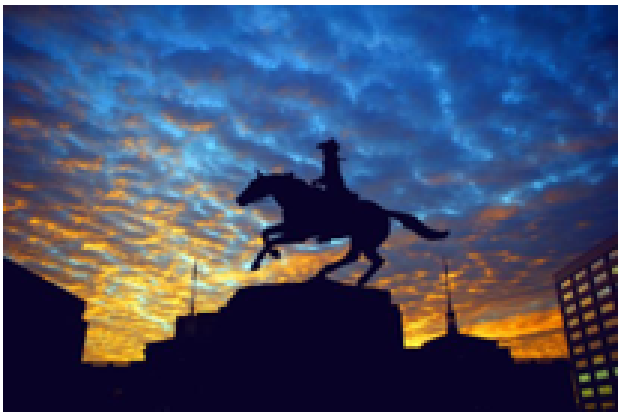
Another example of the dangers of custodian self-collection comes from Nat'l Day Laborer Org. v. U.S. Immigration and Customs Enforcement Agency, government employees from several different agencies attempted to compile data in connection with a FOIA (Freedom of Information Act) request. There was a general lack of supervision and the employees were free to search their own email accounts.

But several ex-employees' emails were never searched at all, which prompted Judge Scheindlin to order the parties to meet again and confer to formulate search criteria and procedures. Scheindlin pointed out that there was no clear identification process in place and the employees were directed to identify personally-created data, which was not a
normal work duty. In fact, one of the major issues was that those government employees were not even aware as to where their custodial data was being stored so, when they merely searched their shared drives, many records were missed entirely.
Despite having spent thousands of hours and hundreds of thousands of dollars on the exercise, the court opined that "Transparency is indeed expensive, but it pales in comparison to the costs to a democracy of operating behind a veil of secrecy". Custodian self-collection has long been disfavored by the courts because it is not systematic, repeatable, and defensible. Judge Scheindlin went so far as to say, "Most custodians
cannot be 'trusted' to run effective searches[…]" This may explain why FRE 902 was amended to require the certification of a collection by a "qualified person". Attorneys should be involved in all aspects of the preservation and collection of their clients' litigation-data, no matter who performs the physical collection. But the new FRE 902 does not directly speak to attorney involvement, which is something we will cover later in this whitepaper.
I would be remiss if I did not mention that almost a decade ago, the Chancery Court of Delaware had all but placed an outright ban on unsupervised self-collection of employee data. In Roffe v. Eagle Rock Energy GP, LP, et al, C.A. No. 5258-VCL (Del. Ch. Apr. 8, 2010), the Court of Chancery addressed party self-selection of the documents to be produced in litigation. "This is not satisfactory. Attorneys should not rely on their client to search their own e-mail system.

[1] Exterro's "Examining E-Discovery Statistics: The Zombie Doctrine of Self-Preservation"

© 2018 doeLEGAL, Inc. All rights reserved.

There needs to be a lawyer who makes sure the collection is done properly."
The State of Delaware has always been at the forefront of the intersection of technology and the law. Widener University (Wilmington, DE) founded The Corporate Counsel Technology Institute to offer the first law school classes on eDiscovery. The Delaware Judiciary created the first court electronic filing system in 1991 and was the first state to have a Supreme Court Commission on Law & Technology (2013).
There was also a historic midnight horse ride by a Delaware lawyer/politician, Caesar Rodney that you may remember from high school social studies class, but that is another topic about firsts, altogether.



Statue of DE Statue of Caesar Rodney's Historic Ride

Where oh where has my data gone?

The over-collection of litigation data is the norm, and we discussed some of the dangers you face in doing so earlier. Collecting enough is not the same as over-collecting. Steve Bunting, CEO of Bunting Digital Forensics, put it this way, "You never know when a spoliation claim will come up and getting everything up front puts you in a far better position down the road".

By creating a forensic image of a device or server, you may just be capturing data that you did not know you would need later. For instance, imaging a laptop will save the entire hard drive, including something called "free space". Without getting too technical, free space is the part of a drive that contains deleted items or file fragments. When someone deletes a file, devices will break that file into several pieces and send it to the free space of the hard drive. A skilled forensic expert may be able to retrieve and reconstruct ESI from this free space which could be important if opposing counsel is accusing your client of intentional or negligent spoliation.

So what's the bottom line here when it comes to self-collection problems? It's really quite simple: most employees don't understand what to collect, or where to collect it from. This is where a forensic collections expert can help solve this problem by creating a "data map" to show all the custodial devices (office, mobile, and personal), any instance of third-party data hosting (emails, social media, sales platforms, etc.), and all the internal and external storage locations. This visual depiction can be very useful because it provides a "data treasure map" that can be easily understood and actioned.

Any good attorney must be aware of the human component and safeguard against risks including the intentional deletion of data and the slothfulness of his/her clients' employees. Attorneys must also have an understanding of the information technology in play. While not every attorney is technologically advanced, having the foresight and ability to retain experts to fill in any knowledge gaps is a desirable attribute--"It's not always what you know. It's who you know."
In every plan there can be challenges and problems. Denis Waitley put it this way, "Expect the best, plan for the worst, and prepare to be surprised." Self-collection of data opens many doors that cannot be closed.

© 2018 doeLEGAL, Inc. All rights reserved.

Top Three Self-Collection Issues

According to a recent interview with Litigation Support professionals, these are the top three areas where self-collection caused issues:

1. A single PDF was created, which contained multiple documents, without any delineation between those documents
2. Metadata had been altered by employees simply forwarding requested emails
3. Emails had been printed and then scanned into PDFs, which removes native metadata

One of those is the perception that a self-collection was not executed properly. By its nature, self-collection lacks any verification of an adequate and complete collection of ESI data. This lack of proof allows opposing counsel to begin questioning the validity of the collection and creating doubt. Doubt is like a Pandora's Box in litigation. Once created, it is an uphill battle to remove. One of the most powerful lessons for attorneys in removing uncertainty and doubt is – document your process. There is no greater position of confidence than knowing you have done everything within your power to be prepared and have a detailed and documented process showing the procedures and tools used to collect the data. We'll get into what this means in a follow-up article.

Technical problems with self- collected data: "The PDF Pitfalls"

Another challenge comes from the resulting changes in ESI data when transferring data from one format to another – most often this happens when creating PDFs from native files. Who're you going to call when these problems arise?

These individuals are the litigation support professionals, the unsung heroes behind every litigation. They fix issues with litigation data and technology, ensuring that document reviews run smoothly. As we mentioned, problems are inevitable, so these professionals often describe themselves as being constantly "in the weeds," "running fire drills," and "underwater." After data is collected, it is sent to these technology warriors to determine how best to make it reviewable.

These legal professionals have seen it all and can reliably forecast common issues with poorly collected and produced data. For example, when a litigation support professional receives a single PDF (Portable Document Format) file containing multiple documents, it presents many obstacles. One problem is that most attorneys will request Bates numbers be applied to each document. Manually separating each document within a single PDF is cumbersome and can lead to mistakes in identifying where one document ends and the next begins.

A PDF is merely an image of a document, and if separate images of each distinct document are not supplied, reviewing that PDF becomes extremely cumbersome. PDFs, by themselves, do not contain a very valuable source of information: native file metadata. Metadata ("data about data") enhances the effectiveness of review tools. It can reveal critical information, such as the type of device the data was created on and date-specific information, which can be very useful at trial. If metadata is altered or removed, the efficiency of eDiscovery technology is significantly compromised.

© 2018 doeLEGAL, Inc. All rights reserved.

Judge Scheindlin ruled that even if a FOIA request does not specify the inclusion of metadata, certain basic metadata information is an integral part of the public record and must be produced. This was the first time a federal judge took this position regarding a FOIA request and the presence of metadata.

She wrote, "Whether or not metadata has been specifically requested — which it should be — production of a collection of static images without any means of permitting the use of electronic search tools is an inappropriate downgrading of the ESI." The absence of metadata creates another issue for review teams: the inability to use de-duplication technology. Most eDiscovery tools have the capability to remove duplicate files through a process known as deduplication. Deduplication typically occurs during the processing stage. However, since PDFs, by themselves, do not contain any native file metadata, deduplication technology cannot function effectively. Identifying duplicate files is done by matching the hash values of two files. Once a document is converted to a PDF, the hash value of the PDF will differ from that of the original file, making duplicates harder to identify. The lack of metadata in PDFs is crippling for several reasons. For one, PDFs do not contain separate text files, so during processing, OCR (Optical Character Recognition) technology must be used. OCR creates a searchable text file by "reading" an image of a document. Although OCR technology has improved, it is not perfect and can produce results that differ from the actual content of the document. The presence of a foreign language or slang can further skew results. When a PDF requires OCR, the same document could appear in the review database multiple times. Much like the children's game of telephone, the same document could be read by different reviewers and treated differently. Disparities in coding could lead to the possible production of privileged documents.

While FRCP 502(d) and clawback agreements protect the producing party from waiving privilege on inadvertently produced documents, the opposing side may discover valuable information within that document, which could be detrimental to your case—"You cannot unring a bell."

One final area all too familiar to litigation support professionals is the practice of collecting data by forwarding emails. While it may seem benign, forwarding emails can alter certain metadata or file folder locations, which might be important. Business users often do not realize they are doing anything wrong, making it crucial to educate all employees on "best practices" and the hazards of collecting data this way. Additionally, when reviewing emails, it is ill-advised to open them in an email application, as this risks forwarding the email or otherwise becoming part of an email chain. All review applications offer options to open a native file in a "near-native format" (e.g., opening an Outlook email file in Word). This allows for defensible review without the dangers of misusing the full functionality of a native application.

Performing a "forensically sound" collection will lighten the burden on litigation support and review teams, as eDiscovery software can function more efficiently. Drinker Biddle's Thomas Lidbury and Michael Boland summed this up in their paper entitled, "Technology: Forensically Sound Collection of ESI," stating, "What makes a collection forensically sound, whatever its scope, is not that the entire storage media has been copied bit by bit, but that the files that have been collected can be shown to be exact copies of what was on the source, including associated metadata." This concept emphasizes using a collection process that will stand up to judicial scrutiny.

© 2018 doeLEGAL, Inc. All rights reserved.

Conversely, allowing non-forensically sound procedures to occur can lead to higher costs, inadvertent disclosures, delayed review times, and late productions. It is always preferable to capture ESI in its "native" format, as this virtually guarantees the metadata is also captured.

This is accomplished by creating a forensically sound copy of the ESI you wish to collect. One way to ensure the defensibility of an ESI collection is to have the attorney involved upfront where they can properly ensure that ESI is not destroyed or altered and all proper procedures are followed. This requires the attorney to implement, transcribe, and supervise a proper legal hold process.

Who is qualified to collect litigation data?

Under the newly amended FRE 902, copied data will not be "self-authenticating" unless a qualified person has inspected the data, recorded the process used, and certified that an exact copy of the data was created. The comment section of 902 states that to meet the inspection requirement, the qualified person can compare the hash values of the original Electronically Stored Information (or ESI) and that of the copied version, but also leaves open the possibility of using other means.

The Advisory Committee on Rules of Evidence has defined a qualified person as one who would be admitted to testify based on their expertise and knowledge of how the data was collected., This definition suggests that the same criteria that one might use to screen a potential expert witness (FRE 702) should also be utilized when deciding who should be allowed to collect litigation data. In regards to data gathering, Judge Scheindlin posited that most employees "cannot be trusted", but whether she was more concerned with their technical qualifications or with their impartiality is up for debate.

Either way, it is crucial to ensure that whoever performs the actual collection of litigation data can withstand opposing counsel's scrutiny, or you risk having the authenticity of the collection called into question. Beyond ease of authentication, data could be inadvertently destroyed or lost during the collection process if the person collecting it is not qualified. This could lead to a drastic increase in litigation expenses, possible sanctions for spoliation, and prolonged litigation.

The goals of the amended Rule 902 are: (1) to remove the need to call an authentication witness, and (2) to give the parties the ability to challenge the legitimacy of collected data before trial. Like most of the Federal Rules that deal with eDiscovery, a skilled attorney can weaponize the authentication requirements of Rule 902 by challenging (or threatening to challenge) any collection when they know that opposing counsel is not using a "qualified person" for collection purposes. Ensuring that an expert leads your preservation and litigation data collection is your best defense and will reduce the likelihood of opposing counsel challenging that collection.

The role of the attorney

While a legal hold is not directly addressed in this paper, it suffers from many of the same issues as a collection, including the identification and supervision that the collection process presents. In short, it is impossible to collect what has not been preserved and it is equally impossible to preserve what has not been identified. With so much at stake, much of the responsibility for preservation and adequate collection practices is being shouldered by the attorney which makes involvement in the process throughout the life cycle of the litigation, crucial to success. Early in the identification process, a litigation attorney must be vigilant of which custodial data ay be involved and where that data is being stored.

© 2018 doeLEGAL, Inc. All rights reserved.

The attorney must determine the breadth, depth, and reach of the litigation and safeguard any data that may fall within that scope. This exercise is not limited to data known to be part of an actual litigation but extends to data that may be part of any future litigations. In effect, the FRCP and the FRE make the attorney the fiduciary of litigation data for current or potential future litigations.

Attorneys involved in litigation often do not know exactly what data or evidence may be relevant, so it is reckless to believe that employees in possession of ESI will be able to make this judgment. The best course of action is to meet with opposing counsel and design a legal hold that will reasonably capture as much relevant ESI as possible. This may include scheduled meetings with identified or potential custodians to ensure compliance with legal holds.

The ABA has expressly recommended that attorneys understand their clients' use of technology (Rule 1.1, Comment 8). Additionally, most U.S. states have made technology competency a part of their Model Rules of Professional Conduct. However, some of the language used by the states is quite vague and speaks to a standard of care akin to the avoidance of negligence.

One final area that should be addressed is the "chain of custody." An attorney should maintain a proper, well-documented ESI chain of custody, starting with the identification and collection of the litigation data and continuing throughout the life of the litigation. Any alteration to the original data that is collected will be traced back to the party that was in control of that data at the time of the alteration. Comparing hash values is a quick and easy way to test if a document has been altered, but this also presumes that metadata has been preserved and collected.

Treat ESI like you would any other piece of trial evidence. It is important to know and understand the life cycle of all trial evidence and to maintain accurate records of your efforts.

Conclusion

Investing money upfront to hire a professional or train internal employees to perform forensically sound collections can stave off future pitfalls associated with the use and review of litigation data. Considering that the most costly aspect of an eDiscovery project is the time spent by attorneys or paralegals reviewing the collected ESI (Electronically Stored Information), this underscores the value of investing in thorough collection practices as a justified cost-saving measure.

There are many moving parts to eDiscovery: human, technical, legal, and ethical. By ensuring that proper legal holds and collection practices are followed, an attorney is best positioned to safeguard their client from possible sanctions for spoliation and reduce the resources spent by litigation support and review professionals. Whatever the amount of money that might be saved by following a DIY model, this will certainly be outweighed by extra resources spent by litigation support professionals, reviewing attorneys, and the need to hire trial experts to authenticate the data collected.

The old English sailing proverb, "a stitch in time saves nine," which loosely means that a timely effort today will prevent more work later, is still applicable to modern-day litigation practices.

About the author: Adam Bowers is an attorney, former business owner, and legal technology practitioner who helps law firms and attorneys navigate the complex world of discovery.

© 2018 doeLEGAL, Inc. All rights reserved.

Four Practical Tips to Help Litigation Attorneys Get Started with eDiscovery

1. Develop a process

Develop a documented process for managing collection procedures, focusing on strategies encompassing human, technical, legal, and ethical aspects, and consistently update it as needed over time.

2. Team training

Develop plans and training to ensure forensically-sound preservation and collection are in place.

3. Stop DIY activity

Stop allowing custodians to self-preserve, self-identify, and/or self-collect the data. Establish a qualified internal resource or contract with an outside provider.

4. Guide Clients in Understanding Cost Implications

The attorney's role is to assist clients in comprehending the complete picture of self-collection costs, which extends beyond monetary aspects. Conducting a cost-benefit analysis will determine whether saving a few dollars initially could result in higher costs later on.

About doeLEGAL

doeLEGAL is built on a promise to provide "Smart data, intelligently delivered." Our software and services help corporate legal departments and law firms efficiently manage operations with up-to-date, insightful data that help teams make confident decisions. We facilitate anytime, anywhere control over cases and costs with advanced management tools and elevated support to generate insights and drive successful outcomes.

To learn more, visit: doelegal.com/ediscovery, speak to a trusted advisor - call 1.302.798.7500, or email info@doelegal.com.

© 2018 doeLEGAL, Inc. All rights reserved.