



eDiscovery Assistant

Drafting a Thoughtful ESI Protocol

A Practical Guide

Table of Contents

3	Introduction	33	Crafting a Process for Search Terms that Works for Your Case
5	When, How and Why to Create an ESI Protocol	38	Negotiating a Privilege Log You Can Live With
12	Key Components of an ESI Protocol	43	Issues to Consider from Collaboration Platforms
20	Form of Production, Why it's Crucial and What to Include	48	Planning for the Production of Social Media
25	Why You Need Metadata in Your Life	52	Top 10 Situations You Can Avoid with a Protocol
29	Manner of Production	57	Conclusion

Introduction

The sheer volume and complexity of ESI (electronically stored information) now requires lawyers and legal professionals managing discovery to plan how it will be handled at the outset of a matter. Simply put, if you don't know what you're doing, you're going to cause problems for your client — whether it's additional costs that could have been avoided, losing out on getting key evidence, or the worst case scenario of getting sanctioned.

99.9% of the evidence you have for a matter today is electronic. In civil litigation, planning for and documenting how you will preserve, collect, produce and review that information has become a must. An ESI protocol is the document that sets out the parties agreement on the how, what, why and when of producing ESI. When provided to a court for review and signature, it becomes a governing document in your case, and one that can serve as a basis for sanctions under FRCP 37(b) if a party fails to comply. It is not a one and done document, it becomes the cornerstone of how you will get the information you are entitled to under the governing rules.

Whether you are new to understanding and leveraging the power of ESI or a seasoned pro, this book will provide practical insights from our team who have drafted hundreds of protocols, seen them challenged, recognized mistakes, addressed new sources of ESI (think Teams, Slack, Threads, etc.), successfully enforced them, and achieved sanctions for failure to comply.

Before we dive in, the most important takeaway from this guide is this — your protocol should fit your case. Creating a template with language on the components discussed here is an excellent starting point. But don't fall into the trap of just using any template you find for every case – you need to carefully consider each and every point that is at issue in each matter separately. Some will overlap, but our experience is that every case has unique challenges that have to be planned for. For example, what is the relationship with opposing counsel? How knowledgeable are they about ESI issues? If the answers are hostile and not very, a protocol is even more important because you need something you can enforce.

The second element to successful discovery is starting early, a topic we cover section 1. Planning your protocol using sections 2 through 8 requires you to think through each of the issues you'll have during discovery and how to address them. The third element is to make sure you think through all the challenges of the sources of ESI that are at issue in your case. Planning to rely on social media? Does one side use Teams or Gmail? Sections 9 and 10 cover the issues in those challenging sources of ESI.

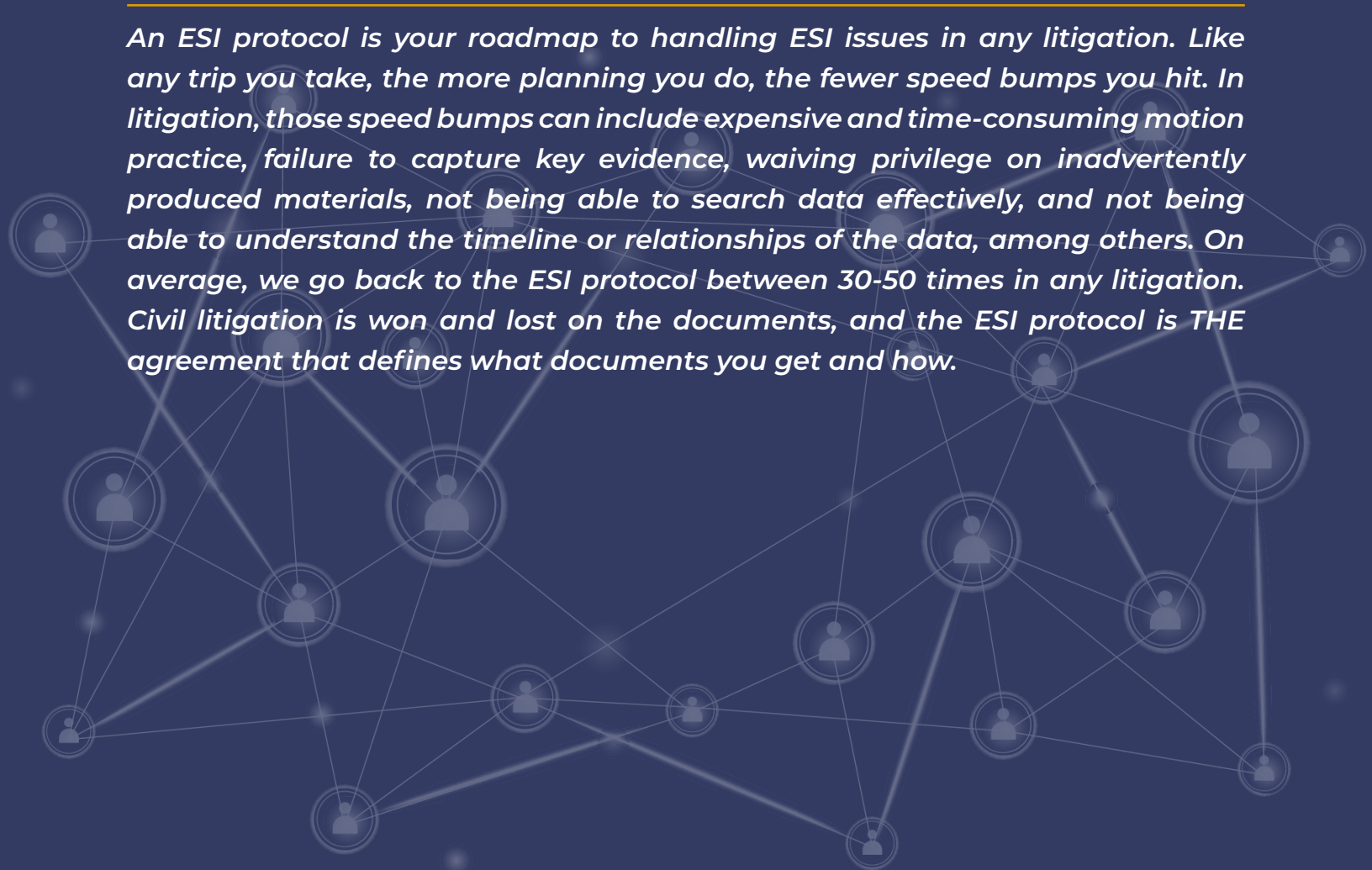
This book is meant to be a practical guide for litigators and legal professionals to work through in creating your protocol. We've provided all of our best tips for what to think about. If you or your team don't have the technical knowledge to work through these issues, find someone to partner with who understands both litigation and the technical issues. You need to think through not only how you'll get data, but what you want to do with it later and plan for that at the outset. The best ones can help you do both.

Good luck.

Chapter 1:

WHEN, HOW AND WHY TO CREATE AN ESI PROTOCOL

An ESI protocol is your roadmap to handling ESI issues in any litigation. Like any trip you take, the more planning you do, the fewer speed bumps you hit. In litigation, those speed bumps can include expensive and time-consuming motion practice, failure to capture key evidence, waiving privilege on inadvertently produced materials, not being able to search data effectively, and not being able to understand the timeline or relationships of the data, among others. On average, we go back to the ESI protocol between 30-50 times in any litigation. Civil litigation is won and lost on the documents, and the ESI protocol is THE agreement that defines what documents you get and how.



SO HOW AND WHEN DO YOU PLAN FOR AN ESI PROTOCOL?

The when is simple: from the minute you start talking about a case. Whoever you have that knows ESI, include them in your initial conversations about the case. You should know many of the answers to the questions raised below right away, and when a user has the ability to delete evidence with a simple click, dropping of a phone, or a system can be auto-deleting data, time is of the essence.

HOW DO YOU START PLANNING?

That's specific to each case and why you need to start right away. Having drafted more than a hundred ESI protocols, there is a very important caveat: You will NOT be able to plan for everything. There will always be unforeseen sources of ESI, unique preservation issues, or some issue that you didn't contemplate and plan for upfront. And that means it's critical to have 1) a well-articulated protocol that tells the court how much you thought about and planned for ESI related issues, and 2) a protocol that includes the requirements for proportionality, and meet and confer obligations that clearly state that the parties will meet and confer to agree on new issues that arise after the initial drafting.



Let's dive into the How. But before we do, let's answer a question that comes up regularly — Do I really need a protocol in every case or for small cases? And the answer is: Do you feel lucky?

You can absolutely simplify a protocol for a smaller case in state court, but going without raises all the risks noted above (waiver, unsearchable formats, etc.), and smaller cases usually don't have the budget to handle those risks. So why take them? You don't have to create a brand new protocol for every case. Instead, create a template that addresses the issues you see in every case, and then modify it for new, case-specific issues or sources of ESI. We've included simpler protocols in the ESI protocol section in eDiscovery Assistant for use in smaller matters. Please note, though, that a sample protocol is just a sample — using one without tailoring it to your case may have malpractice-like results.

The main points to planning for your protocol are listed below. Note that they evolve as your case evolves. Think of these items not as a list to follow, but as points to consider as part of a full picture of what you need to include in your protocol to protect what you need in discovery and how you want to provide it:



Plan and draft your ESI protocol as if you are going all the way through trial and showing electronic evidence to a jury.

That is NOT the same as fully executing on discovery when you know the client's goal is to resolve the case early or your likelihood of success is high to win on a motion. Planning and executing are two different things. What does it mean to plan for trial so early? Social media evidence is a good example, and we are seeing it as a source of ESI in every type of case. You find great evidence on Twitter, Facebook, Instagram, etc. How will you capture it? If you screenshot it, where is the metadata for the post? Who will



authenticate that post at trial? What other evidentiary objections are you likely to face? Each one of those issues has to be considered for every source of ESI you have in a case, and it requires careful planning and a protocol that gets you what you need from the other side. Planning for trial means thinking about admissibility of the evidence, how you will present it, which witness will testify about it, and what form it will take. If you don't get the information in a form that you can use effectively, you will find yourself either 1) spending a lot of time and money to recreate it in a format you can use or 2) not getting the full effect of the evidence at trial. Going back to the social media example, if the opposing party produces a Facebook profile to you that the user downloaded, all of the component parts — images, comments, posts, etc. — are taken apart and produced to you separately. That's how Facebook makes the data available. They do not appear the way they look on Facebook, and you will not be able to re-create them. You need to capture the profile to show a jury the way they are used to seeing Facebook using an eDiscovery specific tool. If you can't get that from the other side, you have to get it.



Next, understand the full scope of evidence that you will want to present on each element of each theory of liability, potential affirmative defenses or counterclaims.

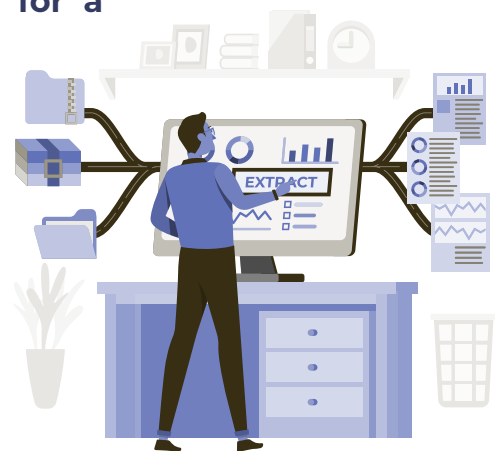
Start with the jury instructions for your case. You want to map out the story you want to tell to a jury and use that map to guide your requests for documents (i.e. ESI) AND then plan for each type of data you will receive in your protocol. We will get more into the specifics on planning for individual sources

in future posts, but this now needs to be a part of your trial strategy. Who can say that the light was red at 10:01pm in the southbound direction when the cars collided? How do I want to show that to the jury? How will I lay the foundation for that evidence? All too often, eDiscovery starts by either 1) sending very broad discovery requests or 2) identifying custodians and responding to individual RFPs. There's very little thought to why you are doing what you are doing, and then you get mired in discovery and lose the big-picture focus. If you define the focus early, you can keep going back to it when the other side requests new data, or you want to make sure you've produced everything you need for your case. Clearly, both sending RFPs and responding have to be done, but planning for the case you want to present and crafting discovery for that process is a much more cost-effective and efficient process.



Determine the picture you want to create for a jury.

This is key. What is the picture you want to create for a fact finder in your case and what ESI do you need to do it? For example, in an action for selling user's data collected from their online activities, the ESI would be completely out of the box, requiring data from databases that has to be combined into a profile for a user to show a judge how a user's data was sold and the path it took. How will you recreate that picture in your case? In an embezzlement case, you want to create a timeline-derived list of all the data and sources taken and where they went. That's piecing together data from multiple places. The more native form you have the data in, the more you can do with it using technology.



Identify the discovery issues that you see in the case. Are there preservation considerations on one or both sides? How far back does the date range for preservation go such that you need to consider the availability of data? Will the other side fight you on getting text messages? List out the issues for each custodian and source of ESI and address them at the meet and confer.



Think through how you envision the discovery process playing out and how you'll plan for it via the protocol. As a plaintiff, can you effectively evaluate search terms and is that the route to go in identifying responsive data? How cooperative is opposing counsel on ESI issues? How will data get identified for review and production? What kind of volumes of data are at issue in the case and what are your resources and budget to review that information? If they are small, technology that allows you to leverage fewer people's resources may be important.



Understand the custodians and data sources you need to account for and what data you want in discovery. ESI is all over the map in terms of how users create, send, receive and maintain it, and then you have to account for how the systems they use maintain data. What kinds of data do you want? What form does that data take? Most lawyers who draft ESI protocols know how to handle email and attachments, but what about social media discussed above? How do you deal with email threading and what metadata fields do you want? Do you know how much you can learn from metadata? What about images, mobile devices (Android v. Apple devices), and instant messaging (iMessage, WhatsApp, Signal, Messenger, Yammer, etc.)? In this step, you need to identify each source of data you may want or need, and plan for how you want to receive the data so you can leverage it effectively.



Identify the date range for preservation of data.

This date range, coupled with the availability of data will focus you on either what needs to be done quickly to ensure data isn't lost, or if data is already gone, whether there are any options available to recover it or fight about it. You need to meet and confer about how far back data is available, and you may need to do some investigation. Start paying attention to the types of data you use in your cases and how long they are available. Bank records? Usually 7 years online. That's just one example to get you thinking. Identifying these issues early can save time and money, and needs to be factored into the evaluation of the likelihood of success on a matter.

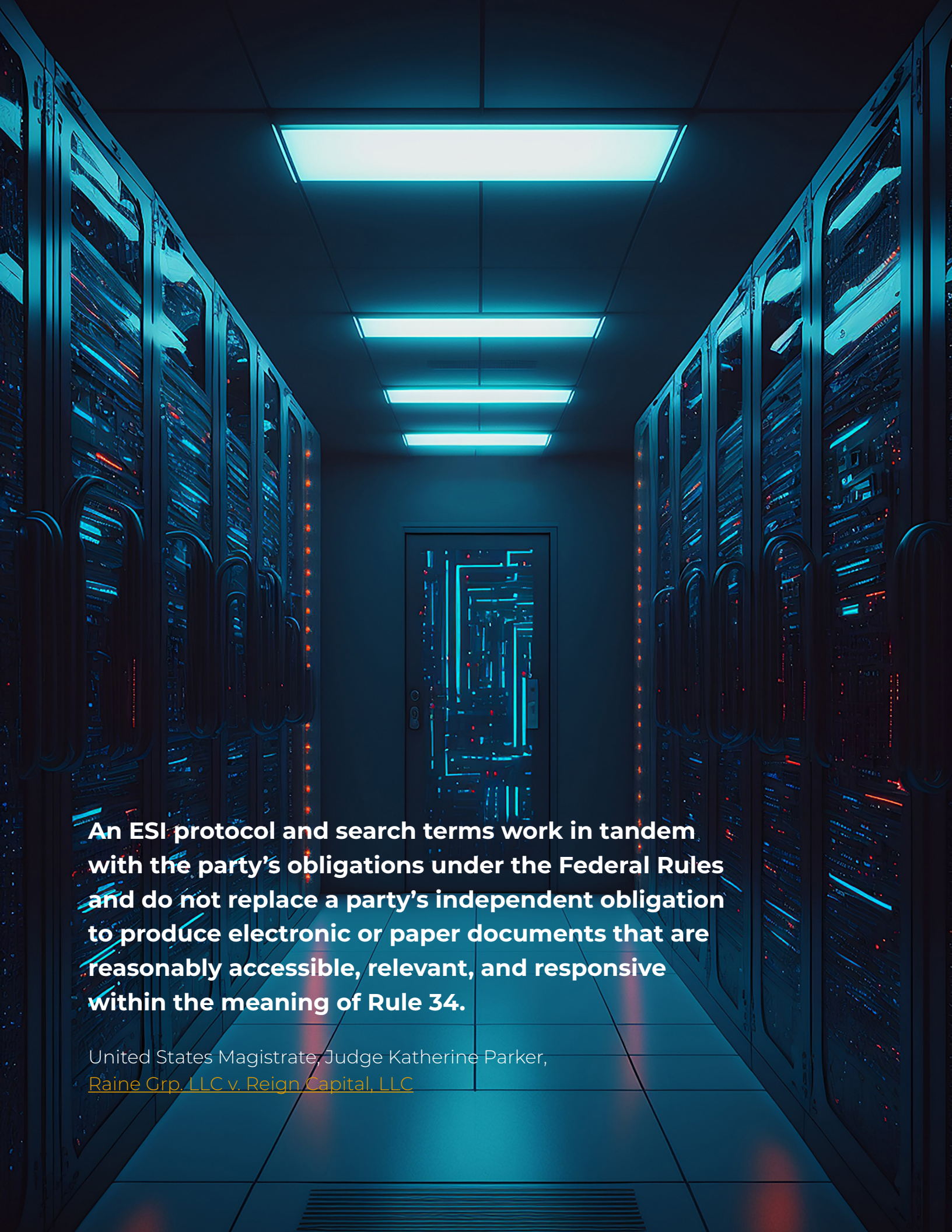




Understand the technology that you will use to collect, review and produce the data and what the other side wants.

FRCP 34 allows the requesting party to request a format, and if there is no request, you can provide any reasonable format. We aren't yet at a place, unfortunately, where we are providing native data, so format is important. In addition to format, you want to consider how the technology you have will cope with the data you want, need or have to provide. Do you have the ability to run analytics, email threading, and filter on the fly? What do you need to be able to do with the data you receive? Because when it comes to the protocol, you want to get the right format for specific types of data and you have to consider how your tech requires it.

The planning part of the ESI protocol is critical to ensuring the issues your client needs for the case are captured in the language of the Agreement. Courts hold parties to what they agree in an ESI protocol. Going forward, we'll talk about examples of how parties gave up rights to contest certain issues based on the protocol. Don't enter into one blindly — consider what you really need. ESI isn't going away and you've got to get up to speed or find someone who is to help if you need to.



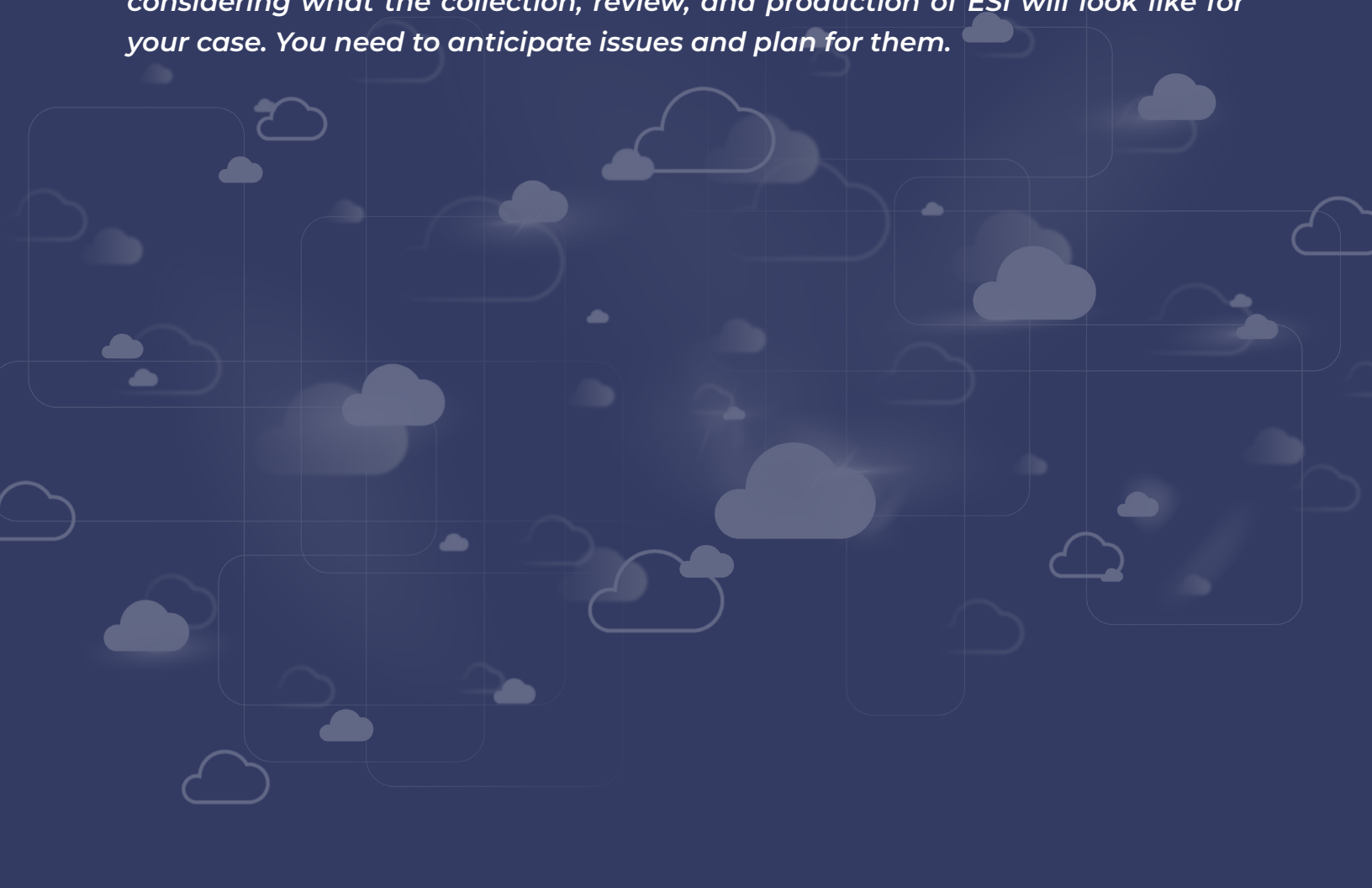
An ESI protocol and search terms work in tandem with the party's obligations under the Federal Rules and do not replace a party's independent obligation to produce electronic or paper documents that are reasonably accessible, relevant, and responsive within the meaning of Rule 34.

United States Magistrate Judge Katherine Parker,
[Raine Grp. LLC v. Reign Capital, LLC](#)

Chapter 2:

KEY COMPONENTS OF AN ESI PROTOCOL

The most important part of crafting an ESI protocol is to make sure it fits your case. Much like a protective order and/or scheduling order requires you to consider the individual needs of a matter, drafting an appropriate ESI protocol requires considering what the collection, review, and production of ESI will look like for your case. You need to anticipate issues and plan for them.

The background of the lower half of the page features a decorative pattern. It consists of a grid of rounded rectangular shapes, some of which are filled with a light blue color, while others are empty. Scattered throughout this grid are various stylized cloud icons in shades of light blue and white. The overall aesthetic is clean and modern, with a focus on geometric and organic shapes.

Some questions to ask before drafting a protocol include:

- **What are the sources of ESI that you anticipate being at issue?** Be specific — email, attachments, mobile devices for text messages, photos, etc., instant messaging platforms, social media, source code, etc.
- **Where are the sources and what has to be done to collect them?**
- **Are the sources all within your custody or control?**
- **Will you have third parties that have data relevant to the case?**
- **What technology do you have to handle data?** Will you handle data internally or hire an outside provider or discovery counsel?
- **What are the data issues specific to each source of ESI you have identified?** i.e. what format will data be provided in, what metadata fields do you need, etc.?
- **Are mobile devices at issue and how will those be captured?** Different methods have different implications.

Generally, drafting an ESI protocol should come AFTER you have conducted custodian interviews and have answers to these questions so you know what needs to be included for your case.

WHAT ARE THE COMPONENTS OF AN ESI PROTOCOL?



Governing Rules and Limited Applicability. Start your protocol with a statement on the rules that govern the case. Include the federal or state rules and the local rules. It's important to define what governs the order in the event of a dispute, and not all states mirror the FRCP. This section should also limit the order to the captioned action and state it does not apply to future litigation.



Cooperation. A short statement indicating that the parties agree to act in good faith and cooperate consistent with the Court's guidelines. This is important when the court you are in has a specific requirement articulated for cooperation, but it really should be included all the time. And you should cooperate.



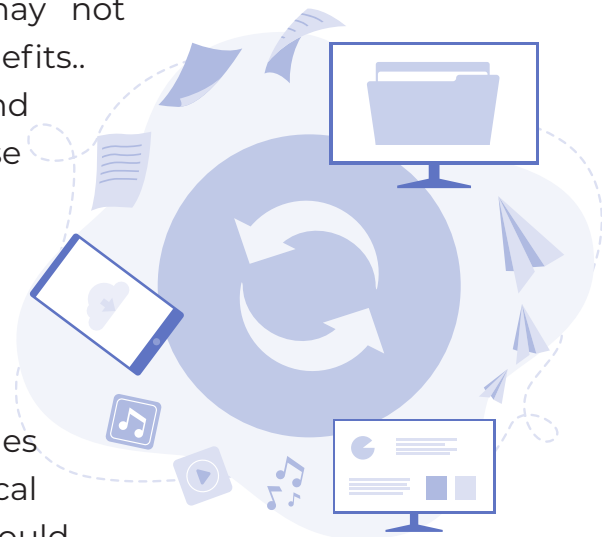
Proportionality. A simple sentence that the parties agree that the ESI provided will be proportional to the needs of the case. While proportionality can be used as a weapon to narrow discovery (particularly in class action matters), courts want to see that the parties know and understand proportionality and will live within its bounds. Note that the revised FRCP incorporate the Rule 26 statements on proportionality into each one of the discovery rules.



Liaison. Several courts now require that each side appoint a liaison for eDiscovery issues, meet and confers, etc. This section should identify those liaisons for the court. This is optional unless required by the court.



Definitions. Your protocol may or may not include this section, but there are benefits.. Should you have motion practice and need to discuss specifics, having these terms defined in the order will assist you in getting that information before the judge. Definitions may include: ESI, native file, metadata, load file, static images, OCR, producing party, receiving party, discovery material, etc. Most judges won't understand them, but the technical folks handling the data will, and you should start to learn them.



Process for identification of Custodians, Exchange of Information and/or ESI Disclosures. This is optional but can be beneficial if one side suddenly asserts you should have provided data from additional custodians or additional data sources down the line (and that happens frequently). The process simply states if the parties agree on a number of custodians (include how many), when that information will be exchanged, what information will be provided about each custodian, whether parties can request additional custodians (yes, on a good faith basis with some justification from data or deposition testimony), and how it should be done. For data sources, it provides a timeframe for parties to exchange information on the data sources for custodians and non-custodial sources (think email systems, databases, etc.)



Scope of Preservation/Preservation Obligations. Whether to include this section depends on several factors that are unique to each case. The model order for the Northern District of California includes a list of sources that are considered inaccessible that the parties do not have to produce, but frankly, that list is very outdated with the technological advancements that have been made, and it often results in fights about whether the receiving party should get those categories of information. The concept of inaccessibility is much narrower now than when that form was drafted. That being said, if you want to limit data sources that will not be searched or produced, this is the section to do it in. This section can also include a date range for preservation that defines the scope of what will be searched.



Search/Use of Technology Assisted Review (TAR).

We are including these together here, but you'll likely want to split them out if you are negotiating a TAR protocol. If the parties are proposing search terms, setting out a process for how those terms will be determined, how they will be applied, whether search term reports will be provided to the opposing party and the process for negotiating terms should be set out. Similarly, if the parties are agreeing on using TAR, that protocol will be its own section or even a stand-alone order on how that will be conducted. Note that Rule 34 does not require the parties to exchange search terms or agree on what technology can be used to satisfy a party's obligations under the FRCP.



Production Format/Form of Production. This is the most complex part of the protocol and one that we'll cover in detail in subsequent parts of our series here. This area can be handled multiple ways. Our experience has been to include all of the details about the production format in one section of the protocol. That includes native v. TIFF and specifications for images, metadata fields to be provided for each type of ESI, extracted text, if native files will be provided for what types of documents, whether and how placeholders will be provided,

applying bates numbers, de-duplication (global vs. other options), de-NISTing, email threading, maintaining families, what is in color vs. black and white, labeling for confidentiality in compliance with a protective order, whether signatures will be separated out as attachments, and redaction. Depending on the complexity of the case, redaction may be a subsection that needs to denote the categories for redaction and the specific wording to be used for redacted material, as well as how metadata fields will be populated to allow the receiving party to know what documents were redacted, and whether native redaction tools can be used during review. Keep in mind that you'll want to have separate paragraphs to deal with unique issues for sources of ESI. If you have source code, instant messaging, text messages, etc., you want to spell out how those must be produced. Consider including that the parties will notify each other of technical difficulties with received productions.



Manner of Production. This is a separate category that many overlook, and it's separate from the form of production.

Form is the file type you receive, manner is the organization of all the files. Although FRCP 34 provides that documents can be produced 1) in the ordinary course of business or 2) in response to requests, we rarely get either with ESI. Information produced by source, by custodian or with the native

file path as a metadata field helps identify the source. However you request data be provided, you are stuck with Rule 34 unless you get agreement from the other side. Manner of production also includes physically sending the productions — whether by SFTP (secured file transfer protocol like Sharefile or Egnyte), encrypted hard drive, etc. Note too whether the parties will send a cover letter or email to identify the production.



Sampling. In certain situations, conducting a sampling of data to determine the value of potential relevant information comes up. If you foresee that in your case, you'll want to include a catch-all paragraph that states that if sampling is required, the parties will meet and confer on a process for identifying the data

to be sampled, how results will be provided, and if possible, the measure of whether the results indicate the identified relevant data is proportional to the burden to get it and what cost-shifting may be available.



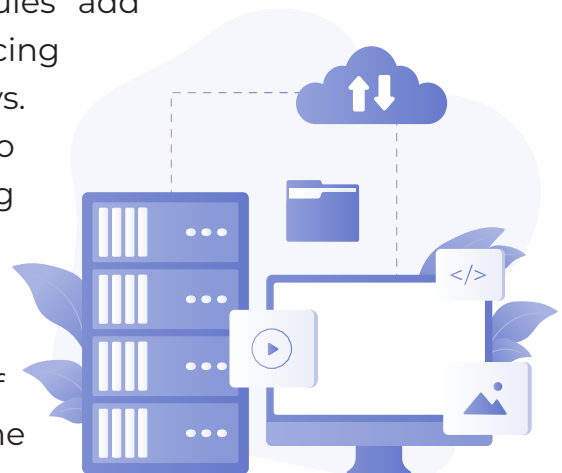
Phasing. In complex cases, it may make sense to agree to a phased approach for discovery. Phasing can include specific custodians, sources or categories of data or date ranges of data. The parties can agree on a phased approach that the court signs off on and allows them to point back to later as necessary. This language can be very helpful if data produced at the end of discovery opens up a new avenue that needs to be explored and you need a basis for asking for more time to explore that limited topic. Courts are more inclined to provide time when it is focused and has a basis for the request vs. that you just didn't get discovery done in time because you dragged your feet.



FRE 502(d) or No Inadvertent Waiver. There are competing arguments as to whether to include this section in the ESI protocol vs. having a stand-alone order. If you use the same language in the protocol and set it out as a separate section header, there's no reason not to include it, but it's a matter of personal preference. Both are orders by the court once signed. FRE 502(d) precludes waiver of privilege for inadvertent production of privileged or protected materials, IF the parties agree to a 502(d) order. You can find language for this section on our sample ESI protocols and sample 502(d) order in eDiscovery Assistant.



Privilege Log. The FRCP set out the requirements for a privilege log and many local court rules add additional information on timing for producing a log (contemporaneous with production vs. at the close of discovery). It's a good idea to incorporate those rules or state how a log will be generated (electronically using metadata is common), how you'll deal with email threads on a log (separate entry for each one or one for the thread), what, if any, additional fields you want added to the log and the process for disputing entries on



the privilege log, and the timing for producing a log – i.e. whether a new log will be issued after each production. Many courts are now requiring logs to be issued within 30 days of a production.



Modification of the Order. Generally, include a paragraph that this order can be modified by the parties or the court for good cause shown and what to call the subsequent order so the court and the parties track which order is currently in play.



Signature Block. This order should be signed by the parties and submitted to the court for signature with the appropriate certification of service.

The list above is an outline for what to include in a protocol. While there are many, many specific details that can be added, an ESI protocol just needs to fit your case. If you're unsure about technical specifications, consult with your provider or discovery counsel to review them. There is no one way to organize a protocol, but some components make sense to group together for readability, and having those first three sections listed here upfront signal to the judge that you know and understand that proportionality and cooperation are key and the parties are on that path. Form of production and the pieces that go into that for specific types of documents are best grouped together so the technical folks know what to follow.

This is a long list, and one that I'm sure seems overwhelming. Many of these pieces may not be at issue in smaller cases, but you need to know what you are deciding to leave out. Take it one piece at a time and consider the value for your case. Build a template for your cases that you can work from that includes just what you need. You can refer back to our checklist on the Components of an ESI Protocol and use the sample orders provided in eDiscovery Assistant as a starting point. We advise against taking those sample orders carte blanche and using them — they have to be adapted for your specific jurisdiction and case.




Notably, file system metadata makes electronic documents more functional because it significantly improves a party's ability to access, search, and sort large numbers of documents efficiently.

United States District Judge Geoffrey Crawford.
[Hoehl Family Found. v. Roberts](#)

Chapter 3:

FORM OF PRODUCTION, WHY IT'S CRUCIAL AND WHAT TO INCLUDE

If you've drafted an ESI protocol, you know that form of production—i.e. the format in which you agree to receive data for each type of ESI at issue—is one of the main reasons that protocols are so key. Absent a written agreement on form of production, or specific instructions in the requests for production, Rule 34 of the Federal Rules of Civil Procedure allows a producing party to produce data in any reasonable format.



What is a reasonable format? That issue is still open. Some courts have held TIFF images and load files are reasonable, others have held non-searchable pdfs are reasonable (yes, I agree with your incredulous expression). Only the most progressive courts have required native production.

You have an opportunity to take Rule 34's reasonable format language out of the equation by drafting a protocol that sets out exactly how the parties will deal with each issue on form of production. There are many issues to consider, and form of production is absolutely NOT a cut-and-paste job from your previous protocol.

As with each section of the protocol that we've discussed in this series, the form of production section needs to 1) address each of the issues on form for each of the types of ESI that you expect to have in your case, 2) to take into account the review platform you will use and 3) consider what you need for how you want to present information at trial. We'll focus on 1 and 2 for this section.

The most crucial part about form of production is to think through exactly how specific sections of your protocol will play out in practice, and make decisions about what you want and how you'll use the data when you receive ESI. If thinking through how you will use ESI at trial is not something you know or understand, talk through what you want to have happen with your discovery counsel, litigation support or service provider and get feedback on how to accomplish what you need. As is always the case in eDiscovery, the more thought you put in upfront, the less problems you'll have later.

Form of production is generally a heading in a protocol, or a separate section, depending on how you set up your document. For each protocol, consider whether and how to include each of the following elements and how they play into what you need for your case and the technical requirements of your eDiscovery review platform:



No downgrading of usability of data. Generally, we provide that ESI will be produced in the form specified in the protocol and that no producing party can reformat, scrub, or alter ESI to intentionally downgrade its usability. That can mean converting data that makes images fuzzier, hides text or alters data in any way that precludes the receiving party from having the same reasonably usable data that the producing party can use.



Specifics on form of production for individual types of data. This can get complex when you have cases involving source code, specific databases, social media, mobile devices, text messages, instant messaging or other cloud-based applications. While we all know what TIFF images plus a load file means, you need to make sure you provide the exact specifications on the size of images, structure of the load file and the other technical specifications that your team needs for the platform you are using. We generally provide different specs based on the source of ESI — think Snapchat, Twitter, WhatsApp, etc.

- » If you are requesting TIFF Images with a load file, you need a separate paragraph that details what types of documents will be produced natively. Typically, this includes PPT, XLS, CSV, Audio, Video, all image files, and non-standard document formats. Placeholders with bates labeled slip sheets are always key when documents are produced natively so that you know when you are viewing the images in document format. The slip sheet will show up telling you that the document was produced natively, so you can click to the native viewer or download the document for viewing.



Family relationships. Generally, you want to preserve parent-child relationships, but you need to also provide how to deal with family members when one member is privileged or needs to be withheld for other reasons like privacy or confidentiality of third parties (hint, put them on privilege log). How to deal with non-responsive documents in families is another issue and one to address in the Form of Production section of the protocol. This issue is now way more complicated with the use of pointers to documents by MS Teams, Slack and Google Mail. Think those through as well.

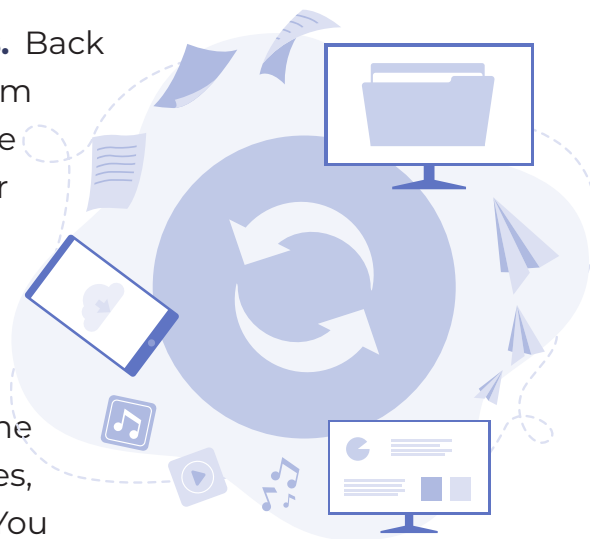


Metadata fields. Most often, we see metadata in chart form that addresses ALL of the metadata fields for each type of ESI and includes the Field order for loading, Field Name, Description of the Field, and an example of how the data

will appear in production. Know your sources of ESI and plan for the specific metadata fields you need for each source.



No obligation to manually code fields. Back in 1997, we printed emails and coded them manually into Summation. Now that we produce metadata, that is no longer needed as a general rule. But it's helpful to state with specificity that fields that are not automatically generated by the processing of ESI or do not exist as metadata need not be manually coded. The exception is the following fields: BegBates, EndBates, BegAttach and EndAttach. You need bates numbers to match up documents, so don't forget that piece.



Extracted Text. This can be a separate paragraph or added to the form for each type of ESI. Wherever you include it, the goal is to provide that the parties shall provide extracted text for all files that originated in electronic form. If extracted text is not available, the producing party should OCR documents and provide OCR text, and OCR text should be provided for any documents that originated in hard copy but are being scanned and produced electronically.



De-NISTing documents. The National Institute of Standards & Technology (NIST) maintains an industry-standard list of file types that are not relevant for eDiscovery purposes. It includes things like operating system files, etc. and is pre-loaded into your review platform such that when data is loaded, it will De-NIST or remove those file types that should be excluded.

- » As a sub-category here, oftentimes parties will exclude additional file types, like signatures from emails, that are junk files (they don't contain any material information). You can agree on the parameters of this limitation and also include language that if a receiving party can show a need for information in those file extensions, it may request them.



Redactions. How you handle redactions depends on what you will need redacted. Is your case one where unrelated, confidential business information may be redacted? If so, you need to provide how a party identifies that information in the redaction box, and whether that information may be logged. Plan for each type of redaction that you may have. The protocol should specify what language should be used in the redaction box for privilege and work product as well.



Bates Numbers and Confidentiality. Address the range for bates nos. and abbreviations for the parties to use in identifying documents and stay consistent throughout the case. We put bates nos. in the lower right corner and confidentiality designations in the lower left. Make sure your system starts at the next number every time you run a production. However you position these items on the document, be sure to articulate it in the protocol so there can be no questions. Make sure to reference your protective order on file, or include the levels of confidentiality in the protocol (confidential, highly confidential, attorneys eyes only). Pay attention to apostrophes and makes sure lit support follows what you include. We've had counsel complain that we used Attorneys' Eyes Only sometimes and Attorneys Eyes Only others. Yes, really.



Deduplication. Decide on what type of deduplication you want—global is the most common using the hash value of the document so that you eliminate exact duplicates. There are several other types of deduplication that you can run once you receive data to expedite review, but global dedupe is the best generally.

Form of production is the key component to getting and using data constructively in litigation. If you want to capitalize on using the newest technology to create efficiencies in review and production, you need data in the right formats. Go through your existing protocols and make sure you incorporate each of these elements in a way that works for ESI in your case. It may be simple, or stunningly complex—it all depends on the case. Now, get drafting. Even if you keep it in your back pocket for your next case.

Chapter 4:

WHY YOU NEED METADATA IN YOUR LIFE

Two issues drive the complexity of dealing with electronically stored information (ESI): the sheer volume of data available and the various technologies on which the data resides. The volume of data means that you need to be able to filter it, sort it in multiple ways and organize it effectively. Linear review of thousands of documents, not to mention hundreds of thousands or millions, just isn't practical anymore. New platforms where data resides mean we have to keep learning what is available from them and how to get it — think Slack, Teams, WhatsApp, TikTok, etc. Those platforms have all kinds of metadata to allow you to filter and use the data, but you have to know it's there, what you can use it for and how to ask for it.

Luckily, the last ten years have been a technological revolution in the legal tech space that provides us with the tools to understand what is in a set of data (whether a production, or data from your client) much more effectively. But to use that technology, you need metadata — the information about the information. Metadata is what allows you to use technology to help you understand what you have in those thousands or millions of documents.

THE FIRST STEP IS GETTING IT.

Generally, a request for metadata and a specific listing of metadata fields to be provided is included in either the Form of Production section of your ESI protocol or as part of an Appendix to the protocol that lays out the specific file types and other issues to be provided as part of the production of ESI. The more complicated the Form of Production, the more often I use an Appendix to set it all out. You can, however, include a request for metadata in the instructions with your Requests for Production if you don't have a protocol in a case.

When requesting metadata there are some key points to make sure are covered:



Provide that the metadata (and any coded) fields that can be extracted will be provided for each document.



Include language that says the parties do not need to manually code any of the metadata fields that cannot be extracted, generally with the exception of Begbates, Endbates, BegAttach, EndAttach and Custodian.

Those fields are not extracted, but generated by the review platform when you run a production — BegBates is the beginning Bates number of a document, BegAttach is the beginning Bates number of an attachment, etc. Those need to be populated as fields and produced.



You can also provide delimiters for the metadata file that allows for your review platform to match the data to the same field in your review platform and allow for easy loading. If you don't know the technical aspects of this, ask your project manager. It varies slightly by review platform.



Reserve the right to request additional metadata fields as necessary in discovery. This comes in handy if you forgot about a source of ESI that has specific metadata fields like social media (URL, account name, etc.)



Include a table with the Field Name and Description of each metadata field that will be provided. Here is an example:

Field Name	Field Description
BEGBATES	Beginning Bates number as stamped on production image
ENDBATES	Ending Bates number as stamped on the production image
BEGATTACH	First production Bates number of the first document in a family
ENDATTACH	Last production Bates number of the last document in a family
CUSTODIAN	Includes the Individual (Custodian) from whom the document originated
HASH	MD5 Hash Value
ATTACHNAME	The file name(s) of the attachment(s) to a parent document(s). Separated by a semicolon

Now I hear your question about the title of this section. Why do I really need metadata in my life? Because you have no idea the wealth of information and relationships you can show between the documents you have using metadata. The AI and concept analytics tools that you know about are based entirely off of metadata (and extracted text). Brainspace, NexLP, Nuix, etc. are all looking at metadata to create relationships between the data that allow you to see how the puzzle pieces that are your documents fit together.

Here are a few examples of how you can use metadata in cases to get to information quickly:

- To get ready for a deposition when you receive a production immediately beforehand, prioritize what you look at by creating a search using Custodian and the To/From/BCC/CC fields for that deponent;

- While reviewing that set of data for a deponent, identify names of attachments or other parties that may have a role, then set up new metadata searches based on what you learn;
- Use date fields to sort documents in date order to review chronologically;
- Tag all the documents you find from the first two sets into a deposition notebook for that deponent, then have them sorted chronologically to create a timeline of the deponent's involvement;
- Use audit log metadata to determine when a party last accessed a system — this is often used in investigating departed employees or potential theft of trade secrets matters;
- Use specific metadata fields to identify whether a group of people fall into the same class;
- Visualize a diagram of interactions between witnesses and dive into how many documents are exchanged between witnesses, what they say and the dates.

The truth is that we use metadata to determine scope for meet and confer negotiations, determining preservation and to answer so many questions that come up in litigation. Didn't preserve a witness? See how many documents are available to/from/cc/bcc to that witness and find out whether you have produced data from that witness, just from a different custodian — e.g. If Mark and Jane talk continuously, and you produce all of Mark's data but not Jane's, it's likely that most of Jane's data has been produced.

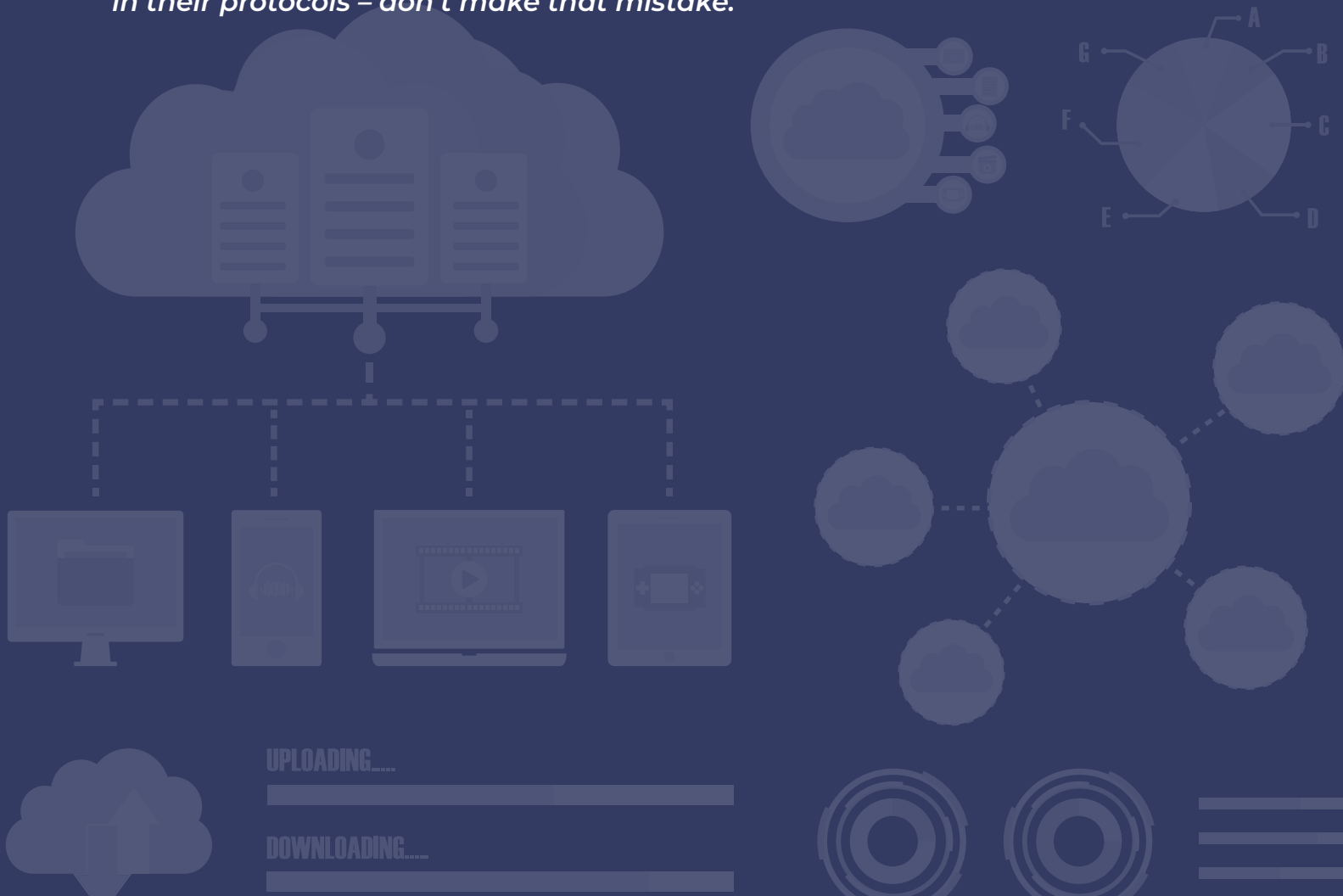
Metadata is a lawyer's happy place when it comes to knowing what information you have in your documents. And since civil cases are won and lost on the documents, it's pretty logical that you want metadata in your life.

eDiscovery Assistant customers can find a [list of metadata fields](#) to request based on the sources of ESI you need included in our ESI Protocols section of Checklists and Forms in eDiscovery Assistant.

Chapter 5:

MANNER OF PRODUCTION

Manner of Production is an aspect of your ESI protocol that directly impacts how fast you'll be able to get data loaded when you receive it, and how effectively you'll be able to organize it for review and use. Most forget about this key piece in their protocols – don't make that mistake.



Manner of Production is HOW you will receive the data from a production. It's separate from Form of Production, i.e. the actual format of the various file types that make up a production. Manner is how the total production will be physically transferred to the receiving party. Options include file-sharing sites or providing physical media. You should NEVER email a production — any leak of data would mean that the hacker had access to your client's entire document production.

Receiving productions of ESI is almost always fraught with complexity — simply trying to access the data, issues loading, corrupt data and other challenges increase the time it takes to get a production loaded. The reality is that it takes up to several hours or days to make a production available when it's received depending on the size and types of the data produced.

Juxtapose the desire to have access to the data immediately with the reality that making a production available is never simple, and you have a recipe for disaster. The lawyers want to see data the minute it is produced; we give new meaning to the term fire drill when a production comes in and depositions are happening days later. We need it NOW. But the reality is that if you don't tell the other side how you want data delivered, you are leaving it up to them and wasting a precious opportunity to get time back on viewing new productions.

If you're nodding along with me, you know the pain of receiving data via SFTP (secure file transfer protocol, e.g. Sharefile) and having to wait hours for it to download, only to discover the data failed to download, files are corrupt, or other technical errors that can result and multiple harshly worded emails are exchanged. Worse yet are failed passwords to decrypt encrypted data, or when someone tries to download using outdated software like Internet Explorer (pro tip: stop doing that and use a different browser). Or when a hard drive arrives and no one has the encryption password because it was sent by letter or email to the lawyer and the project manager is trying to load it at 9pm and can't reach the lawyer.



So, what can you do to avoid these issues and cut down on the time to load productions? Consider the types and volumes of data that you will receive in productions in a matter and include a section on Manner of Production in your ESI Protocol.

What does that mean? Here are some ideas on how to manage Manner of Production that you can incorporate into your ESI protocol as needed:



Discuss the options for transmitting data with the producing party. Many organizations do not have high-speed secure file transfer protocol (SFTP) options and resort to Sharefile or other standard tools. These off-the-shelf tools are designed for small amounts of data — once you start getting into 1GB of data or more, the speed to upload (when the production is complete) and then to download (when the production is received) takes hours on both sides. Consider signing up for your own high-speed SFTP for purposes of the litigation, or use your provider's if you are working with one. Then send an invitation to upload data to the producing party that will cut hours off the upload time, and you'll receive an immediate message when the data is ready for download. One good option that we use is [Media Shuttle](#). Time is money, and in the case of accessing your productions quickly, there is limited time. Get it back by using better technology to transfer data.



Agree that volumes of data over a certain number of GB will be produced on physical media (i.e. hard drive) or via a shared portal that allows for fast downloads.

This goes hand in hand with the first bullet above — if you are using a high-speed SFTP option, you may not need this, but you need to understand the limits of the system you are working with and at what volume it starts to slow down. We will often provide that volumes over 20GB need to be produced on physical media (flash drive, hard drive, etc.) and sent overnight for delivery on the date the production is due. That is key, because many folks will finalize their production the date that it is due, and then overnight it. You lose a day that way. If your discovery schedule is compact and you need documents to prepare for depositions, you'll want to consider this in advance and plan for it.



Know that the available upload and download speeds are affected by your wifi connection speed. If you are working remotely and trying to load a production to a file-sharing site, but your wifi is slower than Sid the Sloth, it will take forever to upload. Consider using an ethernet cable to connect directly rather than using Wifi.



Make sure that your Form of Production is well documented in your ESI protocol so that you can prioritize your data loading.

Less voluminous or problematic data can be loaded separately, meaning that lawyers can get their hands on emails and other unstructured data faster. For example, audio and video files, as well as large images take much longer to load. Similarly, you can load metadata for records and have those available for sorting before the complete catalog of images or other large files might be ready. Lawyers can then tag data that needs to be considered just based on metadata searches.



Document the exchange of passwords for accessing data if they are exchanged — how it will happen, who it will be sent to, etc.

Then ensure that leading up to a production date, those parameters still hold. If a lawyer is on vacation, you need to plan for someone else to send/receive information.



Avoid sending passwords in the same email as the production link.

This is self-explanatory, but there is case law that says if you make a link available in email and your email is hacked, you could be liable for waiving attorney-client privilege in that data.



Ensure the production link expires after a number of days (usually 3 - 7 days).

Another option is to advise opposing counsel that you will remove access once you get notice of a successful download from a shared service.

Manner of Production is often overlooked in ESI protocols. But it can be an important tool in managing the timing of productions and getting access to the data you need, as well as protecting access to your client's data after exchange. Think through what you will need for your case and plan accordingly.

Chapter 6.

CRAFTING A PROCESS FOR SEARCH TERMS THAT WORKS FOR YOUR CASE

With all the talk of artificial intelligence (AI), and acquisitions and inclusions of AI into eDiscovery review platforms, the reality is that search terms are still the first place we start when trying to identify a scope of data for collection. Note the reference to collection, because it's my view that using search terms to identify data for preservation is a slippery slope that should be limited only to situations where 1) you have vast numbers of custodians and data and specific agreements negotiated using the process defined below, or 2) you are responding to a government inquiry that necessitates larger swaths of responsive material than targeted litigation requests should entail. In this section, we will focus on using search terms in responding to discovery in litigation, not for broader investigations.

THE KEY TAKEAWAY REGARDING SEARCH TERMS

How you negotiate the process for deciding on search terms in your ESI protocol requires careful consideration and an understanding of the overall approach to discovery for a case. Failing to adequately consider how this process plays out can leave you and your client stuck with terms that do not cover the gamut of what you need for your case, or with overly broad terms that cost your client too much money to review non-responsive information. Your ESI protocol needs to include a process for identifying search terms, where the search terms will be run (i.e. in what system), what the responding party needs to do to validate those search term strings will provide the most responsive information, and how the parties will iterate on the process going forward.

Think too about HOW you will use this data at trial and how you will want to physically present it to the jury. That impacts the outputs you will want from the sources of ESI at issue. For example, a Celebrite output for text messages is an Excel file. Is that how you want to display a key text string to a jury? Most jurors will struggle absorbing the information when it's visually different than what they are used to, and you'll lose the value of the data.

THE PROCESS

The key to search terms is remembering that the party with the data has the power to know what the best search terms are to be used. The absolute worst way to identify search terms is for a lawyer to sit in his or her office and guess at search terms, or to allow the other side to suggest that the requesting party propose terms first to be tested. The most effective process starts with the requesting party issuing narrowly tailored requests for production, and the responding party proposing search terms on a request-by-request basis. Not a custodian-by-custodian basis, a request-by-request basis. The proposed terms will have to take into account custodians, different sources of data and proportionality. Behind the scenes, the responding party will have already interviewed custodians (or you'll have to get on that immediately) and asked them specifically what terms they use or others use to address the various issues in the case. If the discovery requests have already been issued at the time of interviews, ask the custodians specifically what information exists on a topic, who has it, and what terms would be used to describe it.

Once the responding party has requests, information from custodians and collected data, it should load data and start using the technology to run various iterations of search terms to identify the strings that are most responsive. As you are iterating, keep track of results so that you can explain why /50 is too broad and /10 gives you more responsive results. The responding party provides the proposed strings and the parties meet and confer on whether each one is appropriate or whether further iteration is needed (PRO TIP: It almost always is).

Search Term Hit Reports (often referred to as STR's) are one tool to help iterate on search terms, but they are ONLY a tool. STR's are not the holy grail of what search terms should be for a request or matter generally, and believing that they are, or not understanding what STR's mean, can leave you without critical data for your case. If you don't know how to use them, consider that before including them in your defined process.

When you have agreed upon search results for each request, the responding party runs those results, reviews the data and produces the responsive data. We always ask to have the search terms included as a metadata field in the list of metadata to be provided so that we can filter on the search terms and know what produced documents were responsive to that string. It's an easy filter to do in any eDiscovery review platform (as long as you ask for the metadata field) and lets you analyze whether that search gave you the right documents.

WHAT SYSTEMS ARE BEING SEARCHED

Where the agreed upon terms will be run is the next piece to focus on because there is an enormous disparity in results between systems. For example, the same search string run against multiple custodians in MS Teams will give you substantially different results than pulling the custodian's mail into a review platform and running the search there. Think carefully about where the search terms will be applied and include it in the protocol. It varies by system and the type of data.

Next, consider the different types of ESI that you'll be searching. You wouldn't use the same search terms for email that you want for text messages or instant messaging. Frankly, search terms for text messages are fraught and I suggest steering clear of them — instead consider text strings among custodians and date ranges as a way to filter texts. Custodians will use entirely different language in instant messaging applications (think Slack, Teams, WhatsApp, etc.) than they use in email or texts. Have

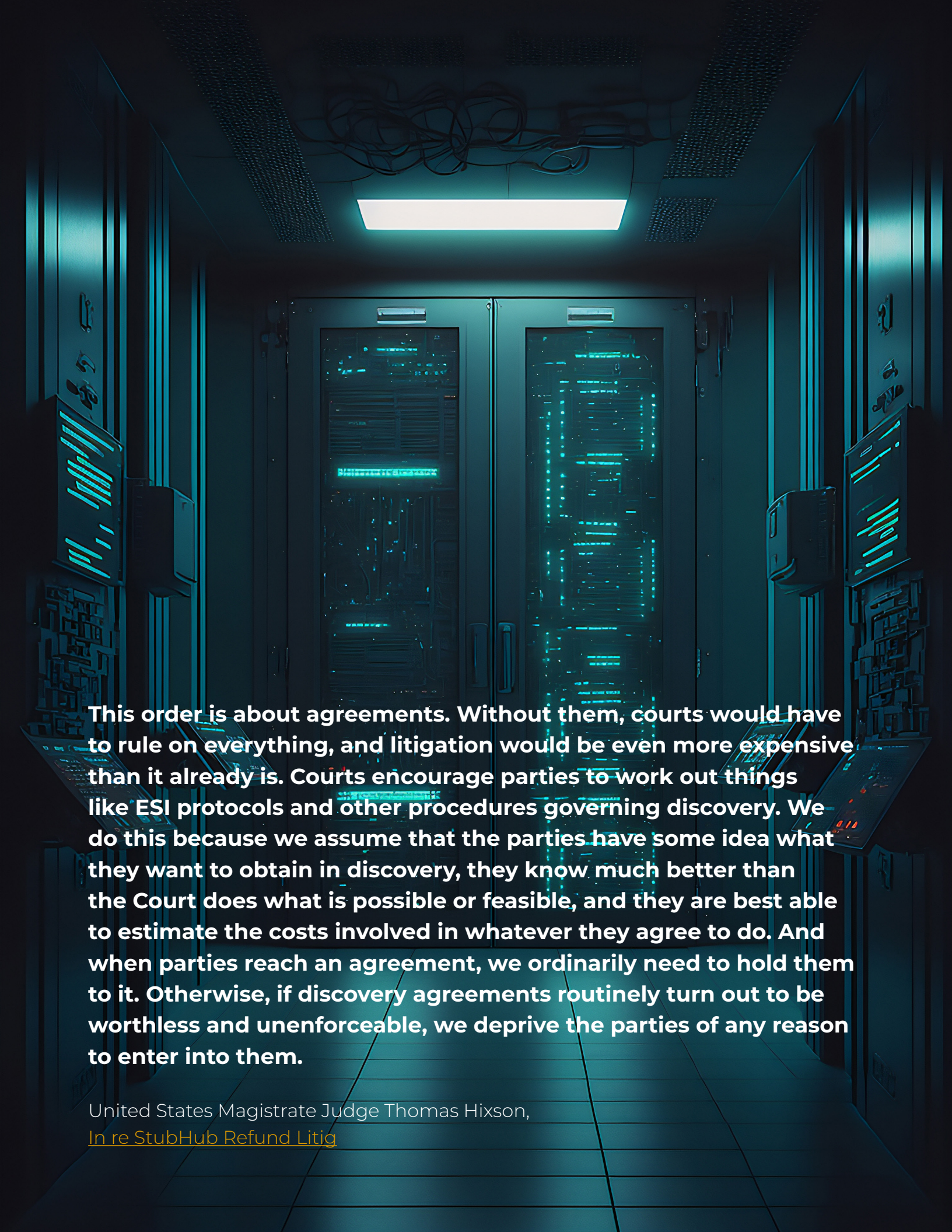
the responding party review the data and propose a methodology that's reasonable — usually channels (for Slack) dedicated to the project with some carve outs for other non-responsive business information, or chats for WhatsApp with specific custodians. There are inexpensive ways to collect this data via searchable pdf that allow you to OCR the data and make it available for search if necessary. We prefer to review these types of data separately, so think about that as you are drafting the protocol.

ITERATIVE PROCESS

Keep in mind that the process is iterative. Once you receive data, review it immediately to understand whether you need to ask for additional terms, additional custodians, etc. based on your review. And make sure your protocol includes language contemplating an iterative process. This is not a one-and-done process, but many producing parties will argue it is once search terms are agreed on initially, and you need to protect your client from that argument.

PULLING IT ALL TOGETHER

Search terms and how to leverage them continue to be a key aspect of the eDiscovery process. To use search and search terms effectively, you need the overall picture of the case and what you will want to do with your evidence in front of a jury to guide each step of the process. This isn't simple, and it requires a great deal of thought. If you don't think you have what it takes to undertake it for your client, find someone who does to assist. Missing out on key search terms can mean the difference between getting the documents you need or losing your case altogether.



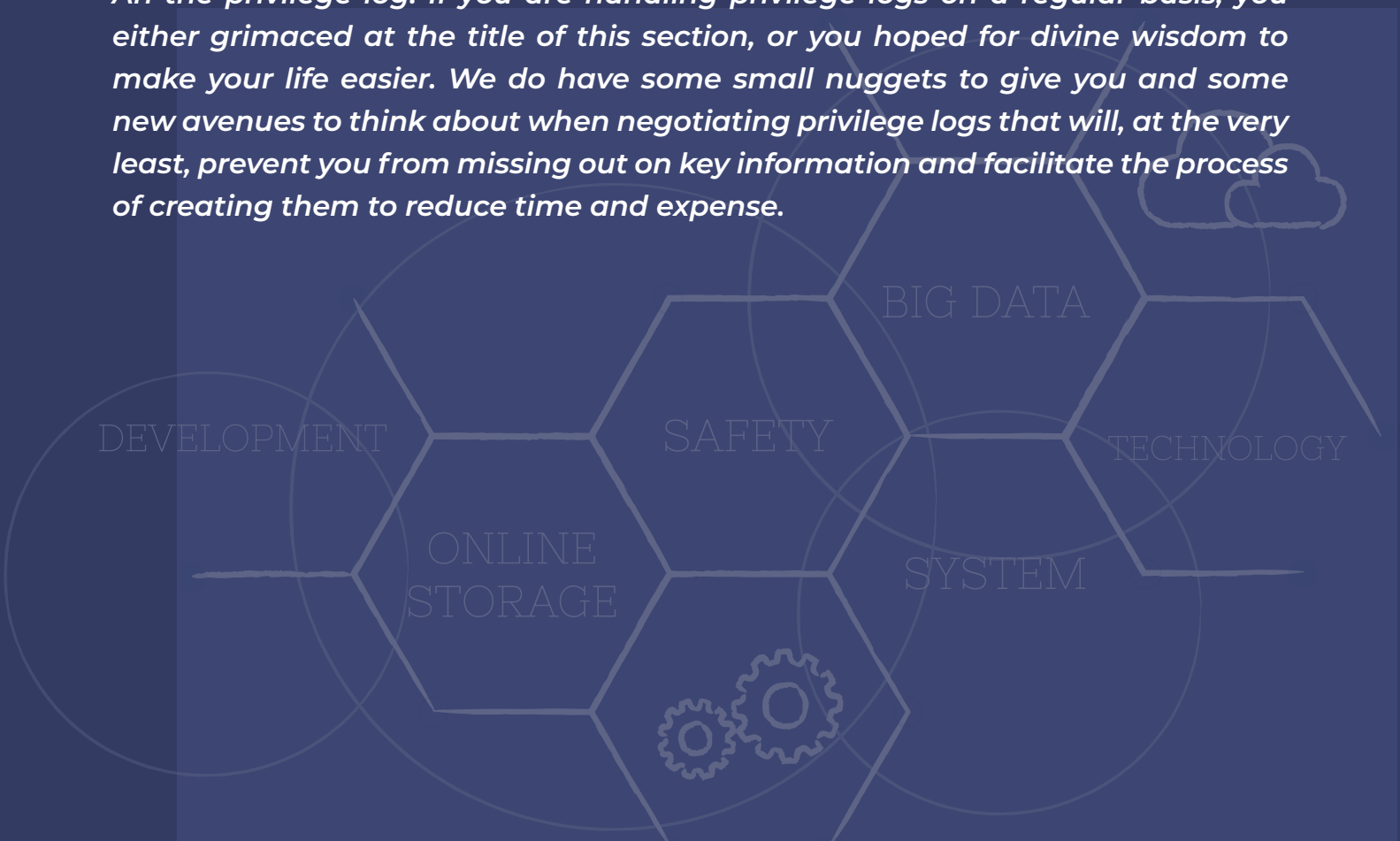
This order is about agreements. Without them, courts would have to rule on everything, and litigation would be even more expensive than it already is. Courts encourage parties to work out things like ESI protocols and other procedures governing discovery. We do this because we assume that the parties have some idea what they want to obtain in discovery, they know much better than the Court does what is possible or feasible, and they are best able to estimate the costs involved in whatever they agree to do. And when parties reach an agreement, we ordinarily need to hold them to it. Otherwise, if discovery agreements routinely turn out to be worthless and unenforceable, we deprive the parties of any reason to enter into them.

United States Magistrate Judge Thomas Hixson,
[In re StubHub Refund Litig](#)

Chapter 7:

NEGOTIATING A PRIVILEGE LOG YOU CAN LIVE WITH

Ah the privilege log. If you are handling privilege logs on a regular basis, you either grimaced at the title of this section, or you hoped for divine wisdom to make your life easier. We do have some small nuggets to give you and some new avenues to think about when negotiating privilege logs that will, at the very least, prevent you from missing out on key information and facilitate the process of creating them to reduce time and expense.



If privilege logs are something that rarely enter your mind, I encourage you to read this post and consider whether you may need to increase the importance of them for your cases.

Now, does the concept of the privilege log NEED to be included in the ESI Protocol? The easy answer is no, it doesn't, but it makes sense to include it there since the protocol is governing the production of ESI, and what I suggest below means that you can add the information needed for a privilege log during first pass review if you think through and set it up correctly. It also keeps all the parameters about the production of ESI in one place.

FRCP 26(b)(5) governs the requirements of a privilege log:

(5) Claiming Privilege Or Protecting Trial-Preparation Materials.

(A) Information Withheld. When a party withholds information otherwise discoverable by claiming that the information is privileged or subject to protection as trial-preparation material, the party must:

- i. expressly make the claim; and
- ii. describe the nature of the documents, communications, or tangible things not produced or disclosed—and do so in a manner that, without revealing information itself privileged or protected, will enable other parties to assess the claim.

In addition to Rule 26 (or its state equivalent), many judges have individual or local orders that dictate both the content of and timing when privilege logs must be provided. See the list below for more info on timing.

Most of the time, we happily go along with Rule 26 and any persuasive court authority on what needs to be included in a log. But the advent of ESI for the last 15 years means that there is significantly more data, and as a result, there are also significantly more privilege-related issues that need to be considered. Quite frankly, the rules need to reconsider being more explicit about what is required in a log to avoid some of the unnecessary skirmishes on this issue.

Now, some of those privilege-related issues are limited to more sophisticated commercial litigation, but in reality, privilege issues in almost every matter, whether for plaintiff or defense, small or large, complex or simple. Long story short, the technology that we use to review and produce documents allows us to create an

electronic privilege log that can be produced with the push of a few buttons, and there's no reason not to provide it or ask for it.

For example, collecting a user's inbox at any organization in which that user may have received emails to or from:

- in-house counsel on unrelated matters,
- in-house counsel related to the matter, but that are business advice and not legal advice
- third-party providers who are implementing legal advice but may or may not have privileged protection in the jurisdiction where the matter is pending
- outside counsel on unrelated matters,
- outside counsel on the pending matter PRIOR to the filing of the complaint, or
- outside counsel on the pending matter AFTER the filing of the complaint.

We've also been seeing more and more that corporate individuals are using text messages, WhatsApp and other platforms to exchange messages, so each platform will require its own search protocol for identifying privileged information. For example, when a platform only has phone numbers, you'll need to have the phone numbers of counsel to ensure those are captured.

There are also multiple considerations on the types of privileges and circumstances where waiver can come into play that's beyond the scope of this section, but that you'll want to consider.

SO, HOW DO YOU ACCOUNT FOR PRIVILEGE ISSUES IN A PRIVILEGE LOG WITHIN AN ESI PROTOCOL?

The first step is to consider all the practical issues associated with the log and how the facts of your case may impact the following:

- 1. Date Range of Entries to be Recorded.** Typically, parties do not require the producing party to include attorney-client privileged communications after the date of the filing of the complaint. Most of the time in litigation, the con-

duct that spurred the claims is in the past, and while you will certainly have communications that are privileged discussing the matter after the date of filing, parties typically agree that privileged communications post-complaint are not logged. That can change if the facts that are relevant to your case are ongoing, or if there's a separate reason to require them.

2. What Should be Included in Log. As noted with the language of Rule 26 above, the courts require enough information to allow the receiving party to understand the basis for the claim. What that means for each case depends on the sources of ESI at issue. Most parties interpret that language to provide the author and recipients, date sent, summary of the content of the document, and the basis for asserting privilege, as well as any additional information required by the court. The summary of the content can often be where the time and expense of creating a log comes into play. Emails are easier in theory — they have a subject line that can be listed on a log. But what about when the 5th email on string of ten that has been marked privileged and has nothing to do with the original subject line? That's when you need to require a breakdown of what is privileged and why. Consider using these fields for your team to complete during first pass review on your eDiscovery platform, then link them together in the final Excel spreadsheet as one column:

- Description of Privileged Substance — information that describes the content
- Description of Privilege Name — this is who generated or received the request for legal advice
- Privilege Description — a short description of who is requesting information from whom and about what
- Privilege Basis — Attorney-Client Privilege or Attorney Work Product
- Privilege Substance Notes — Optional and notes for the producing party's use about what the content is
- Privilege Substance Regarding — Optional and notes about what the privileged material is regarding, for example “sales information, payroll records, email accounts” etc.

3. Timing of Delivering Privilege Logs. The big issue here is whether the parties will exchange a privilege log within a set timeframe after each production (say 30 days), or if the parties agree to provide one complete log of all documents withheld for privilege at the close of discovery. That may turn on the Judge's order— more and more judges are requiring a log for each production within a reasonable time—or it may turn on how long discovery is and whether the parties have similar discovery obligations. When asymmetrical litigation is at play, things get more complicated.

Planning for identifying privileged information and creating a log needs to happen when you first start a matter. We create privilege filters for clients and update them regularly. That's step one. Step two is thinking through what you will need from a privilege log from the other side, and what you will provide, as well as what unique considerations you may need in terms of fields for specific types of ESI you know will come up in the case. Step 3 is negotiating the scope of the log — whether in your ESI protocol or elsewhere. Step 4 is to plan your review to create the fields you need for your privilege log and have dedicated reviewers for a second pass privilege review to ensure it's ready to go upon production or soon thereafter.

Including the privilege log scope and production considerations in your ESI protocol allows the parties to understand the full scope of what is required, and has the judge sign off on it. That makes enforcing it easier and erases any doubts about the parties' obligations.

Planning for privilege issues starts before litigation. Know your clients, create filters, understand the privilege issues that recur over and over again so you can anticipate them, know the law on privilege and meet your obligations to create and produce a log that complies with your obligations.

Chapter 8:

ISSUES TO CONSIDER FROM COLLABORATION PLATFORMS

The pandemic vastly accelerated the use of collaboration platforms as organizations scrambled to have ways for their employees to stay connected and communicate in real-time. Meetings normally conducted in person were relegated to online video meetings via Zoom or Teams and companies that had not yet adopted these tools rolled them out very quickly to allow business to continue while we sheltered in our homes. Tools like Teams, Google Apps (now [Google Workspace](#)), Slack and Zoom suddenly became commonplace.



Fast forward to the post-pandemic era, and these tools are now the cornerstone of workplace communications. Email is still a primary source of external communication — to folks outside the enterprise — but internally, employees lean on the chat functions and other built-in tools of these collaboration platforms. As with any new tools, that means there are eDiscovery implications to consider. These new collaboration platforms include instant messaging, chat, file sharing, collaboration rooms, video conferencing and online meetings, VOIP (Voice over Internet protocol), phone systems and apps galore for folks to use.

If you've made it this far in this guide, you already know what comes now — the complexity of these collaboration tools means you have to plan for dealing with the preservation, collection and production of a whole new range of types of ESI that will come from these sources. We touched on the need for planning in Chapter 1, and collaboration tools make that planning process even more important. There are several unique considerations with collaboration platforms that need to be considered early on, or you could end up 1) not getting the information you need in the way you want it, or 2) having to redo discovery at great cost.

HANDLING POINTERS OR LINKS TO DOCUMENTS

The biggest issue for eDiscovery professionals right now is the change in the way these collaboration platforms handle what we used to call attachments, and the parent-child relationship. Previously, emails had documents physically attached to them so you always got the document attached to an email in a production (or you should have), and it was the version that the author of the email had attached at the time. New systems — Slack, Teams, Gmail, Dropbox, Box, and multiple other tools now provide links to electronic files that either remains static, or more likely is updated regularly. So the question becomes when the document is NOT physically attached to the email or chat message, how do we preserve, collect and produce those documents to maintain the relationship created by the link? You need to plan for that in your ESI protocol.

This issue is so complex that we haven't even agreed on what to call these links or pointers to a document. In [Nichols v. Noom](#), the Court called them hyperlinks and stated that hyperlinks are not attachments. That was controversial, but not wrong — not all hyperlinks are attachments. Some are just links to something on the internet and we've never considered those attachments in the past.

There's some notion that "pointers" is the more appropriate term, as the links "point" to the document where it lives. A quick search of the term "pointers" shows that [a pointer is an object in computer programming](#) that stores a memory address. A pointer is really just a set of characters that "points" to the location of a document, photo, or other type of file that the user wants to reference.

This issue of pointers, or links, has multiple implications for your ESI protocol:

- You need to understand whether any of the sources of ESI that your client or the opposing party has include pointers or links.
- You have to be able to communicate in technical terms with opposing counsel (and they with you) about how to handle producing documents that are at the other end of the pointer/link while maintaining the parent-child relationship, if possible.
- You need to include specific language on how each party will handle responsive ESI that appears under links included in responsive email, chats, etc. Your protocol needs to define terms, determine how data will be collected, and then produced to keep those relationships intact. The language and specifics will depend on the platform you are using, but it needs to be agreed upon upfront. Courts, including Judge Parker in Noom, have shown little sympathy for parties who do not understand the data issues up front and will not order a party to redo discovery because you didn't realize you needed to have linked documents produced with a parent-child relationship.

VERSION CONTROL

The other significant issue with the technology of using pointers or links is which version of the document is produced with the corresponding message. Yes, you read that correctly. A pointer or link to a document that you sent six months ago points to the same place, but the document that lives there may have been updated ten times since you sent that link. Standard eDiscovery protocol is that the version that existed at the time the pointer or link was sent should be the one that is produced with that message. After all, what the author of the responsive message intended to attach reflects their communication most accurately.

Most of the collaboration platforms store each version with a time stamp, so it IS possible to get the version sent at the time of the message, but it involves thinking about that before any work is done to identify the scope of data and it's complicated depending on the platform. It is crucial to coordinate with your discovery counsel or project team to ensure that you are considering versions when preserving and collecting data for production.

FORMAT AND LOCATION

We covered form of production extensively in Chapter 5, and collaboration platforms add even more complexity to what you want in produced documents. For example, Slack data is exported in .JSON format, and needs to be converted to be readable. That means you need a tool that is designed to export and review Slack data that allows for searching. Even the smallest Slack collections contain thousands of messages, and you'll need to be able to filter them with search terms or TAR, both of which work best with native data.

But format of the ESI at the pointers or links is also critical. The most obvious choice is native format with metadata that allows a link back to the message or document that pointed to it. But our tools in eDiscovery have not yet caught up to handling this issue, so it's crucial to understand what you can do, and what you are required to do BEFORE you get started.

Location is also tricky. Teams data is dispersed across the platform and requires accessing it from multiple locations. Private chats, group chats, uploaded files and pointers or links are all in separate locations that need to be tracked down.

Both Slack and Teams allow for connecting hundreds of other applications as well, meaning that you may have to go to third-party applications to collect ESI that would have been "attached" in earlier days. Those issues significantly ramp up the complexity of a collection and require much more manual intervention that we got used to with collecting email and its attachments.

PULLING IT TOGETHER

Collaboration platforms have made work easier — no question. But the process of finding, keeping, collecting and producing data from these platforms has introduced challenges that the technology in eDiscovery is still grappling with. Know what your challenges are, and be sure to choose professionals and tools to ensure that you have a defensible position in handling these types of ESI.



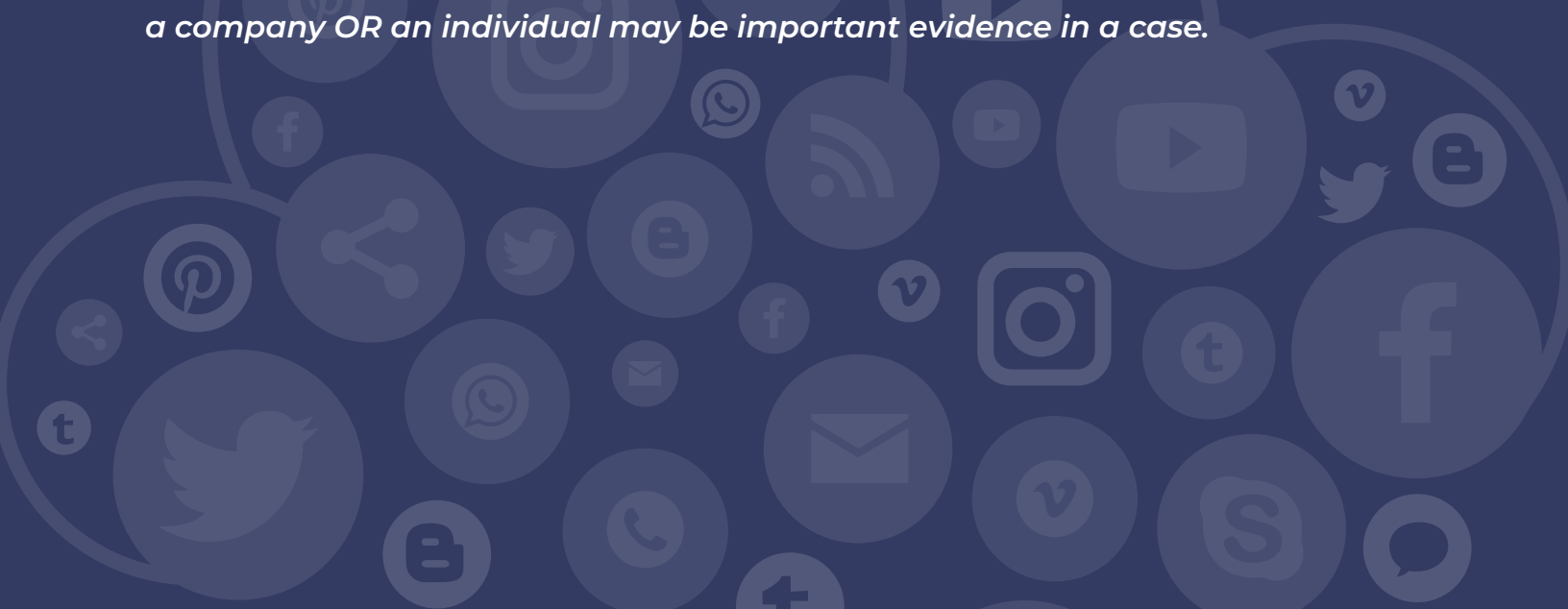
Let's get back to basics. Litigants should figure out what they are able to do before they enter into an agreement to do something. Litigants should live up to their agreements, especially when they are embodied in court orders, as the ESI protocol is here. And if for some reason a party learns that a so ordered discovery agreement has become impossible to comply with, the party should promptly move for relief with a good showing that despite its best efforts, compliance is impossible.

United States Magistrate Judge Thomas Hixson,
[In re StubHub Refund Litig](#)

Chapter 9:

PLANNING FOR THE PRODUCTION OF SOCIAL MEDIA

Drafting an ESI protocol that meets the needs of your case means identifying each type of data that is likely to be produced in a case and including the parameters for that data type in the protocol. Lately, we've seen a big issue with parties failing to provide specifications in protocols for a key type of data that's showing up regularly: Social Media. If you are an employment lawyer, dealing with marketing issues in commercial litigation, or working on trademark infringement or other IP claims, social media evidence is key. But we see it as well in family matters, class actions, and any place where the online presence of a company OR an individual may be important evidence in a case.



So, how do you plan for social media evidence in an ESI protocol? Consistent with other data types, your protocol needs to include language to deal with the big 3 — the image, text (if it exists) and metadata. But social media adds another wrinkle, and that is the information about a social media post — think likes, comments, and other means to react to a post that differs by platform. Will those reactions be included in what you capture? Are they necessary for the case? That will vary case by case, and you need to consider it during collection and what you want to include in your protocol.

First, a bit of terminology. An individual entry on social media — namely a post on Facebook or Instagram, a tweet on Twitter — is called a post/tweet/etc. Contrast that with a user’s entire profile that contains all of the posts, tweets, etc. You need to understand whether you need/want individual posts or the full profile.

Here’s a list of several steps you need to work through to include the appropriate language in your ESI protocol to ensure you get the data you need about social media:



Identify the list of social media platforms or posts you want. To do this step, you can’t just guess, you’ll have to get on the computer and start looking for what social platforms contain data you want produced in discovery. Each platform has its own unique features and metadata that have to be considered — what you need from Tik Tok is not the same as what you need from Twitter, Reddit, Instagram or Facebook. Note that we are strictly talking social media here. Issues with instant messaging (including WhatsApp, iMessage and others) and other sources of ESI are covered in later sections of this guide. What platforms is your client on that need to be preserved? What platforms is the opposing party on that you want to have preserved/produced? Make a list of what you need, the date ranges covered or the subset of posts, and the metadata that you want from each platform. More about metadata two steps down, but you want to start thinking about it when you identify the platforms.



Identify the scope of data you want from each platform. Make a list of the date range(s) covered by the case or the subset of posts (i.e. a topic or set of topics that are covered) that are relevant. Spend time on each platform that is publicly available to really drill down and know whether you can create a list of URLs by post using a topic or keywords that you can search for, or if the posts are so voluminous that you want to preserve the entire profile. YouTube videos

are one-offs, you're unlikely to collect the whole channel. But a Facebook profile dedicated to a brand may be entirely relevant over a specific date range. Be thoughtful for each platform and get on and see how the parameters you ask for will affect the scope of the data you get. Couple that with the bullet below on how you want to show evidence at trial. All of these work together.



Identify the metadata you want for each platform. Each platform has different fields of metadata, and you need to identify them for each. Some sample metadata fields include the URL for the post itself (helpful for Twitter, YouTube, Facebook and others), the Title of the post (useful for Twitter, not as much for other platforms), the SMA Account (Social Media Account) that identifies the author of the post. For each field of metadata you request, make sure the other side can provide it, and make sure you code your review platform with those fields to populate when the data is loaded from the production. Think of which metadata fields by how you will want to sort the data — do you want to sort by date? By keyword? By author of the post? That tells you which metadata fields you need to have. Many of the standard metadata fields you use in an ESI protocol will also provide data on the social media posts including Date, Custodian, RecordType (be sure to specify this will be social media), TimeSent, etc. Each of the recommended metadata fields for documents and social media are included in the List of Metadata Fields in the ESI Protocol section of the Checklists and Forms on eDiscovery Assistant.



Consider how you will want to show posts or profiles to the judge or jury at trial and then request image files. Often, we ask for the preservation of entire profiles for social media, but we often want to either have individual posts as exhibits, or we want to create a montage of posts that show the behavior at issue for trial. Think carefully about what you will want to show and the context

you want it in when drafting your protocol. Often, we ask for specific posts individually with all metadata and then a full profile to show context of the individual posts.



Review the posts when they are collected or produced to make sure your images and comments appear the way they do on the sites so a jury will recognize the content. QC (or quality control) upon collection or production is key here. Don't just load data and forget about it, you need to make sure you have what you need, that the images are the correct size, they are of sufficient quality to view them and be blown up as exhibits. Verify that you have all the metadata fields for each post, test the URL to make sure it links properly, and verify the image looks good. You'll be in a world of hurt if you don't have an image at the time you have to compile trial exhibits and the post no longer exists.



Forget about using the Facebook download of a user profile. Facebook created a setting that allows a Facebook user to download their profile, but the download separates all of the file types so that the data appears nothing like it looks on Facebook. You'll have files of images, text, and links, but none of it will be related, and comments, likes and other data aren't captured in that download. And it's almost all in HTML, which isn't read by most eDiscovery platforms. In short, it's rarely a good preservation or collection tactic.

Adding another layer of complication, around the time of the 2020 Presidential election, social media platforms altered their APIs to prevent election interference. That had the added issue of complicating the collection of social media by current tools and new privacy considerations are ramping up those changes. Do tests when collecting, and make sure you can provide the metadata you are both asking for the other side to provide and that it's available. Work through these issues BEFORE you agree to a protocol, and you'll have a much easier time using your social media data when it comes to trial.

Social media is crucial evidence, and to get all of it that allows you to filter, sort and present it, you need to plan for it in your ESI protocol. Do the work upfront to reap the benefits at trial.

Chapter 10:

TOP 10 SITUATIONS YOU CAN AVOID WITH A PROTOCOL

You don't know a good thing until it's gone. Hindsight is 20/20. If only...

There are many different sayings to express regret in not planning properly for producing or receiving ESI in litigation. The one, surefire way to avoid that regret is to get a good ESI protocol in place that considers all of the aspects of ESI issues in a case.

But maybe you're not convinced, or you don't know the real value of why you need to spend time drafting one, and you're not even sure you know what to include in it. We've heard from in-house lawyers time and time again that their uber-sophisticated large firms aren't using ESI protocols in a complex IP or commercial disputes.

WHY NOT? COULD IT BE THAT THEY DON'T GET THE VALUE?

So, as we reflected on all the issues that we see with data we receive, whether subject to an ESI protocol or not, or from a third party who doesn't feel the need to comply with a protocol that is in place. And what showed up time and time again is that to know WHY you need a protocol, you need to understand WHAT you are missing out on for your client when you don't have one. In collaboration with [Joy Murao](#), CEO of our partner, [Practice Aligned Resources](#) in Los Angeles, we've identified ten situations that we regularly see in receiving data, how those could have been prevented with a solid ESI protocol to enforce with the court, and what it costs your client each time one of these things happens.

As you read this list, give thought to the data you are currently trying to wrangle for depositions, what you want to show at trial, or what data you need to prove your case. Notice that I said data, not documents. Data about the document (think metadata about a social media post—who authored it, when it was published, who commented, what they said) is sometimes as, or more important than the document itself.

Here it is, the list of the top ten situations you can avoid with a good ESI protocol:

- 10. You need to search the text of the documents, but you receive unsearchable pdfs or a production with no text.** Simply put, no text = no search. Your protocol has to provide for the production of a text file with images. (Of course, if you're receiving native data, that isn't an issue, but if you are, you likely don't need to worry about this situation.) Can you OCR the documents to create text? Sure you can, but you'll have to pay someone to do it, either by the hour for PM time, or by the document. You also slow down your ability to start reviewing the documents because you can't sort them effectively until you can search across them. OCR is never as good as text from the source, so you'll also have errors.

- 9. You receive inconsistent production formats.** This one happens ALL THE TIME. Even with a protocol, parties will produce documents one way (think TIFF images and a load file) and another way the next time (non-searchable pdfs). They'll then argue that you have a reasonably useable format and you don't think the court will grant much relief, but it seriously impedes your ability to use the documents effectively (see no.1). We've seen productions from third parties in HTML format—no images, no nothing else—that are forwarded from counsel months later and there's no time to go back to the third party. Those documents are barely usable. Your protocol needs to have a production format section that addresses the format for each type of document (think bulk categories and exceptions, like spreadsheets and PowerPoints, social media, images, etc.) that allows you to take a party to court to enforce the production format. Lack of consistency in format adds thousands of dollars in costs to your case and will keep you from being able to leverage data. You can't see the formulas in a spreadsheet if the spreadsheet is an image or in pdf format.
- 8. You receive data with a created date newer than the modified date.** This is a doozy and one we see a lot. You need to ensure metadata is not altered during collection so you receive the same data that the other side has. Your protocol needs to state that the created date is a metadata field to be provided and that data will be collected in a way that preserves the metadata intact. Too many times data is dragged and dropped into a folder for collection, and that process changes the created date to the date it's moved. You've just spoliated data. The fix is to require the other side to re-collect, re-process and re-produce the data a second time. That sounds like fun and something the other side will immediately agree to, right? You can protect your client with language in your protocol that lets you go to the court as needed, or that ensures the other side will do it right the first time.
- 7. Your production includes images or social media posts, but you receive only images with no metadata.** What you have are just images, and images can't be OCR'd. That means you have to manually identify and code text or metadata about each one to be able to use them effectively in sorting, adding them to trial exhibit lists, etc. Instead, request specific file types and metadata fields for both images and social media (they are different) and discuss how they are collected. For images, be sure to get the native file path so you can know where the image came from. If it's not attached to an email

or other document, you want to know what else could live where that image lived in case you want it.

- 6. You need to redact documents or you want to identify redacted documents.** Redaction for reasons other than privilege is becoming much more of an issue as documents showing irrelevant business information are implicated in discovery. I usually want to see all the redacted documents, but unless I asked for a metadata field populated for redactions, the only way to find them is to go through them one by one.
- 5. You need to sort by the Confidentiality designations.** This goes hand in hand with the redactions. I'm pretty keen to see what documents are marked "Highly Confidential—Attorneys' Eyes Only" and make sure that all confidential documents should actually BE confidential. Your protocol needs to include metadata for confidentiality designations and the right language from the protective order populating that field.
- 4. You need to know the original file path of the document.** You find the smoking gun document in your collection, but you have no metadata—no custodian, file path or other identifying data—to know where it came from. We use file path all the time to track where a document lived, and that allows us to know what witnesses to use the document with during depositions. Include it as a metadata field in your protocol. You'll be able to use it in so many ways.
- 3. You need a consistent process and timing to produce privilege logs.** FRCP 26 provides what needs to be included in a privilege log but not format or timing (although some judges do have local rules on privilege logs). Articulating the fields and timing of the privilege log gives you a written order to point to when you haven't received a log, or when you doubt whether privilege applies to certain documents.
- 2. You need a court order for a basis for sanctions.** FRCP 37(b) requires a party to show a violation of a court order to get sanctions under that section and guess what, the judge signed off on the Order you added to your ESI protocol. Viola. You just saved yourself a motion to compel and moved right to sanctions for failure to abide by it.

- 1. You inadvertently produce information.** Yes, the number one reason you need an ESI protocol with an FRE 502(d) order or a sufficient clawback is when you inadvertently produce privileged information. With the fire drill way discovery usually happens, this very situation has been addressed by a specific rule of evidence. Use it to protect your client and incorporate it into a protocol.

This list could be much longer than ten—in fact, we came up with dozens of scenarios that allow you to control costs and your ability to use data better for your cases by having a solid protocol that anticipates what you need and ensures that all parties provide it. Remember that not every protocol has to be complicated, but since every case includes ESI, every case should have some protocol.

Conclusion

Since you've made it this far, you know we weren't joking at the complexity of creating a thoughtful, intelligent protocol for your matter. A reminder of the three keys to success in drafting – 1) develop a protocol that fits discovery and theories of liability for the case you are working on, 2) start early — as soon as you know about or reasonably anticipate litigation, and 3) think through all the challenges of the sources of ESI that are at issue in your case.

Done well, your ESI protocol, working in tandem with a protective order and an effective discovery strategy, can be an incredible tool for managing costs in discovery and ensure that you get what you need with enough time to use the documents and ESI effectively in litigation.

Save this book on your computer or print a copy to keep on your desk. Highlight the parts that you know you need to focus on and read through it each time you have to start a new matter and need to think through these issues. Use it as a guide to everything you need to think about when drafting a protocol. You and your clients will be better for it.

We'll be updating it as new issues arise that need to be considered.

Thank you.



Kelly Twigger
Founder
eDiscovery Assistant

About eDiscovery Assistant

eDiscovery Assistant was founded in 2016 by renowned litigator and eDiscovery expert Kelly Twigger. Trusted by AMLAW 100 law firms, Fortune 100 corporations, government agencies, and legal service providers, eDiscovery Assistant is the only research platform designed exclusively for ediscovery. Our mission at eDiscovery Assistant is to help lawyers and legal professionals leverage the power of ESI for their clients. The platform offers access to a curated collection of over 30,000 eDiscovery case law decisions (with new cases added daily), along with useful checklists, forms, and resources. Organized for easy navigation, eDiscovery Assistant empowers legal professionals with the practical knowledge to tackle any eDiscovery challenge. Reimagine your approach to legal research and unlock new possibilities in leveraging ESI today with eDiscovery Assistant.

TO CHECK IT OUT FOR YOURSELF:

[Schedule a demo](#)



[Sign up for a free 14-day trial](#)



CONNECT WITH US ON SOCIAL MEDIA:



eDiscovery Assistant

Contact

eDiscovery Assistant LLC
2945 Julliard St.,
Boulder, Colorado 80305

 www.ediscoveryassistant.com

 info@ediscoveryassistant.com

 [/company/ediscovery-assistant™/](https://www.linkedin.com/company/ediscovery-assistant/)