

How to Handle Cases with More Discovery than You Can Read

Supplemental Materials for the 2019 Federal Criminal Defense Practice Seminar
for the Eastern District of North Carolina

October 4, 2019

Russell M. Aoki, Aoki Law PLLC
russ@aokilaw.com

Christina Nguyen, Aoki Law PLLC
christina@aokilaw.com

The landscape of our practice has changed. Big cases are no longer measured by the number of banker boxes but by the gigabyte and, increasingly, the terabyte. The search across dozens of boxes for the smoking gun letter is now a search through millions of emails, text, and social media messages. Communications are made up of hastily drafted words that are a mix of letters, numbers, and images. A conversation could start on Facebook, move to text messages, and end with an agreement confirmed by an exchange of emojis. The commonly used metaphor of a needle in a haystack no longer accurately reflects the challenges we face. The haystack is now an open field, and the needle is often broken into fragments.

What is our haystack really made of? Discovery often includes materials clients create using their smartphones, tablets, laptops, and desktop computers. Accessing cloud-based services such as social media significantly increases the discovery type and volume. Ever noticed Facebook subpoenas ask for records from 2004? The Government doesn't know when a client opened a Facebook account, so they subpoena records from when Facebook started its online service. The production is rarely limited to the indictment period but spans the life of the account, which means tens of thousands of pages of posts, comments, photos, audio, video, and private messages.

Law enforcement also makes discovery available in the form of wiretap interceptions, body wire recordings, body camera video and audio, GPS tracker data, surveillance video, lab reports, investigative reports, search warrant pleadings, and video interviews. Much of it is not useful to your case yet is provided by the Government in the "abundance of caution" to ensure they meet their *Brady* and *Giglio* obligations. We've all experienced "over-gathering" by law enforcement. It used to be months of just pole camera video but now it's a lot of everything.

Grand jury subpoenas result in faint image bank records that include fill-in-the-blank checks, emails from web-based services like Gmail, cell tower records, and medical records or maybe mortgage documents.

Too Much Discovery: Technology Companies Saw a Need

You're not the first one to complain about too much discovery. A multi-billion dollar litigation support industry exists to help lawyers manage volumes of discovery so great that no lawyer, even with dozens of paralegals, could review every item. These companies create and maintain powerful web-based databases that use high-speed servers that sit in secure data centers across the country. These platforms offer powerful search engines that can read across millions of pages of documents of varying formats then allow you to categorize results by subject matters you choose, like people, investigative reports, events, or indictment counts.

Just because you're using technology doesn't mean you must abandon the review strategies you've honed over your career. Litigation support tools are built to support the review strategies of experienced litigators like you. The difference is the extra steps of carving down all the discovery productions to subject matter specific collections before you apply your review strategy. Big discovery cases require a two-step process that first addresses discovery organization and second, discovery review.

Some Basic Assumptions Before We Begin

We will address how to use cloud-based e-Discovery review platforms, referred to as web-based databases throughout the memo, that host more discovery than anyone can read. We will also look at example strategies to search, review, and eventually categorize key materials for more detailed analysis. This memo assumes several things have taken place. First, you met with the Government and discussed when and how discovery will be produced. The Recommendations for Electronically Stored Information (ESI) Discovery Production in Federal Criminal Cases (ESI protocols) published in 2012 provide excellent guidance on what to consider and comes with a checklist to review during your discussions. If your jurisdiction does not provide for a "meet and confer," it will soon. Congress is about to enact Criminal Rule 16.1 that will make it a requirement.

Rule 16.1. Pretrial Discovery Conference; Request for Court Action

(a) Discovery Conference. No later than 14 days after the arraignment, the attorney for the government and the defendant's attorney must confer and try to agree on a timetable and procedures for pretrial disclosure under Rule 16.

....

Second, this memo assumes you sampled the materials to see if they are complete, searchable (text and Bates numbers), and there are no corrupt files. You've also identified duplicates, and in a very general sense, anything you need not review, like commercial music saved on a mobile device. This step is worth your time. Although everyone is eager to search, if your documents are not searchable, you'll never know what you're not finding. Separating files you don't care about will quicken your review because your searches will not pull up irrelevant items that take time to open, confirm

you don't care, then move to the next file to review. For forensic computer images, it would mean removing program and system files created to operate computers and exact duplicates in temp files - a process known as "deNISTing" and "deduping".

The third assumption is you have selected a suitable database company to host all your discovery. This is also a time-intensive task because it'll require you to learn how databases work, their search and review features, and costs to make sure you'll be getting market price. A good rule of thumb is to get quotes from at least three database companies. You'll learn a lot from talking to each vendor and will better understand costs. If a Defender Services' National Coordinating Discovery Attorney (CDA) has been appointed to assist you, then don't worry about these three steps. The CDAs have identified cost-effective programs.

Why Trying to Read It All May Not be a Good Idea

What are the benefits of a web-based database? You'll save time and do a better job. A recent study compared linear page-by-page review of documents against computer keyword searches.¹ The test collection was 9,764 documents. Three experienced lawyers did page-by-page review. It turns out they missed nearly 12% of the relevant documents and took 98 hours to complete the task. In contrast, one lawyer and a tech-savvy assistant used keyword searches and found all but 4.28% of the relevant documents in about 14 hours. <http://reasonablediscovery.com/wp-content/uploads/2016/06/ILAS-v.-Linear-Review.pdf>.

The problem with linear page-by-page review is we can't stay attentive long enough. When Marvel's 3-hour movie *Endgame* hit the theaters, many people were hesitant to see it because they didn't think they could sit still for that long. If you were one of those people, imagine someone telling you they wanted you to be part of a 3-person document review project that will take 98 hours!

For this memo, we'll use illustrations from two databases: Invendica and Casepoint. Invendica is often used for cases with a lot of audio and video. It can stream recordings, so you don't have to download audio interceptions or surveillance videos to review them. It's easy to use but does not have the computer analytics to speed up the review of hundreds of thousands of pages of documents or months' worth of emails.

For big document cases that have mostly documents and emails, programs like Casepoint provide powerful search engines and time-saving analytics. Casepoint is a web-based database that contracts with Defender Services, which means if you use it, Defender Services will pay the cost, not the Court. It can do many things with little effort on your part, like find and group email communications. Casepoint also has a predictive coding feature that will help identify big buckets of materials by subject matter in a sea of documents using artificial intelligence. Casepoint takes more time to learn but the more you know its capabilities, the faster and more accurately you'll find things.

¹ *Iterative Legal Analysis v. Linear Document Review. – A Comparative Study*, (Anne Kershaw, Columbia University, June 2016).

Build a Library

After checking the productions, sampling the contents and loading the materials onto a database, the next step is organization. General organization of discovery is much like building a library. Libraries are more than buildings full of books. In its most basic layout, a library categorizes books by subject, with travel books in the travel section, cookbooks in the cooking section, biographies in the biography section. Likewise, categorizing discovery lets you quickly return to each department in your library every time you need to conduct more research.

If your productions are already broken down by subject matter or document type, you may not need to do anything more, or at least little more for now. If you receive discovery with all the search warrant pleadings in one folder, witness interviews in folders marked for each witness, reports marked in their own folder, then you have a good foundation for your database.

There is a benefit to having the database company load the discovery with the same folder names as the original productions. If the Government identifies a document as an exhibit from a specific folder, you can go right to the folder and search within it.

Adopt a Review Strategy

Once on the database, you can create more folders specific to how you want to defend the case. By knowing how you plan on using the database first, you'll be better able to organize the discovery. Here are some basic thoughts:

- Identify broad subjects you want to research and the search terms to use (like all items that mention your client)
- Search, sample, refine your search, sample again
- Save your results as general subject matter categories
- Search within your general subject matter categories for more specific items using the search, sample, search method
- Then review your results and rank the results based on importance (Rank 1, Rank 2, Rank 3)
- Review all your Rank 1 items as you would for any case, and identify which items go to your defense, particular witnesses, overt acts, or counts in an indictment

It can be scary thinking about what you may miss if you don't apply a specific enough search. However, the strategy for how to find everything is the opposite. Starting with broad searches will capture more than you want but lessens the chance of missing something. Sampling your documents and identifying more specific search terms is key.

Example: Start with your Client

A good place to start is looking for every discovery item that mentions your client. If your client is Bill Brown, a search for Bill will come up with more than Bill Brown; that's too broad of a search. It'll come up with other people named Bill and "bill" as it is used for invoicing. You realize that after sampling 25 or more documents. A search for just Brown may lessen the hits but the color brown will come up. A better fix is to add Bill's last name to your search, at least initially until you can sample more documents to find more search terms to add. The search of "Bill w/3 Brown" ("w/3" being within 3 words) will capture pleadings, reports, but maybe not wiretaps (we'll look at wiretaps separately). By adding "w/3", if a middle initial or middle name is present, the search will capture the Bill and Brown on either side.

The process of search, sample, search could take several iterations. As you sample, you'll find more search terms, like instances of the last name spelled Browne or William used instead of Bill. As you sample, you may find related people, phone numbers, and email addresses – things you want to note for later searches.

A subject matter category of your client would have all the investigative reports, surveillance reports, and search warrant pleadings. Client names are frequently misspelled in reports. Clients could have multiple nicknames spelled in different ways. If you want to search for discovery items that mention your client's name, with each search you must include all your client's name variations. However, by creating a category of search results for the different ways your client is identified, you'll only need to run those variations once, create a category, then go back to that category and perform more specific searches within it.

Let's say your case involves your client Bill and his brother Jerry. You can create a subject matter category for materials that capture Bill Brown, using multiple searches for his name then again for William Brown, Will Brown, Bill Browne, William Browne, Will Browne, Willie, Willy, and WB. It may bring 500,000 documents down to 5,000. Now you can search within that category for more specific items.

Search within your Category to Learn about the Investigation

Being early in the case, you may want to understand the scope of the agents' work by reviewing the investigative reports and search warrant pleadings. Search within the category you created with search terms like:

- FD-302
- Affidavit
- Title III or T-III
- Search warrant
- Incident report

- Arrest report

You can save your results as another category like “Reports Bill Brown”. You can search within it again by case agent name, witnesses, co-defendants, or by dates.

During your sampling, Bill’s brother Jerry is often mentioned. You can create another category to find all items that mention Jerry Brown, Gerry Brown, Jerald Brown, Gerald Brown, Jerry Browne, Gerry Browne, Crazy Jerry, Crazy Gerry, JB, and GB. You can find all the investigative reports as you did for Bill.

Now, you can search within one or both categories at once to find every instance where “gun” is also mentioned and save your results. If weeks later you realize the gun is also referred to as “Glock”, you can go back and search the word “Glock” within the Bill and Jerry Brown categories and save what you find in your “gun” category.

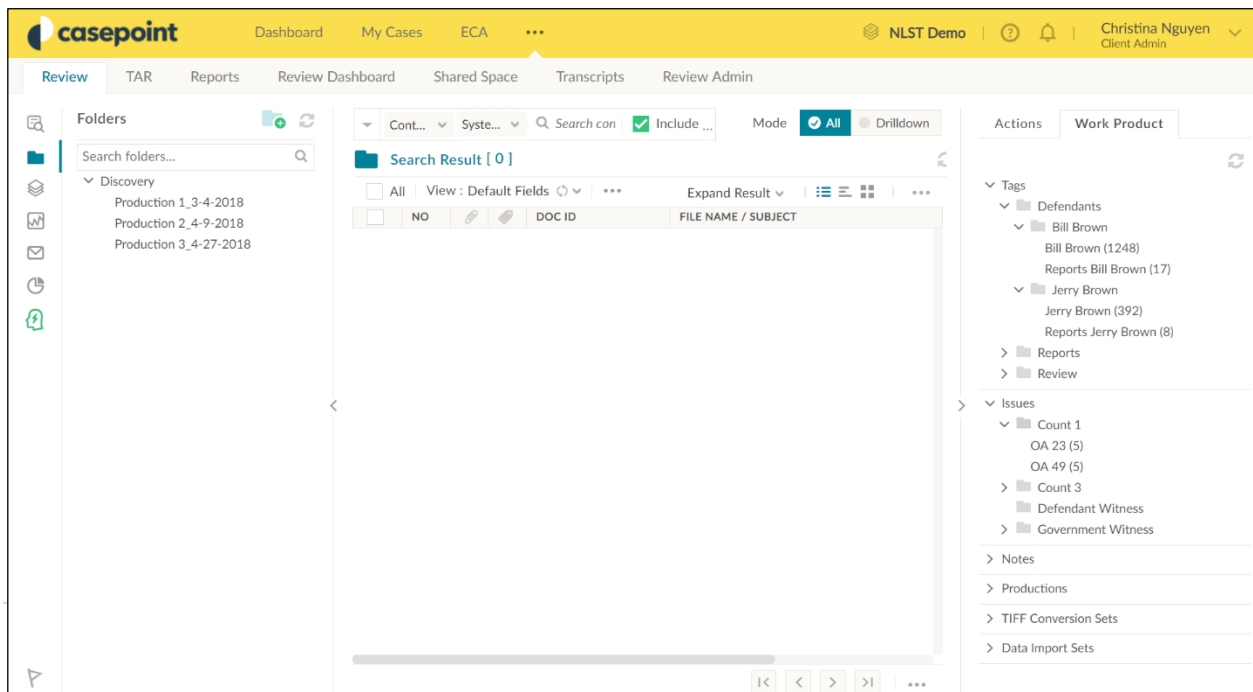
Imagine manually doing the same thing for 500,000 documents or even 5,000 documents without a technology tool. How long would it take you to manually find and collect the same materials? Is it likely you’ll miss something?

Using Casepoint to Search Documents: This is What it’ll Look Like

Let’s add pictures to what we’re talking about to better envision categorization.

Databases differ in how they make subject matter categories. When we’re working on our desktop and laptops, they are referred to as folders. Casepoint refers to their work product folders as “Tags” and “Issues”.

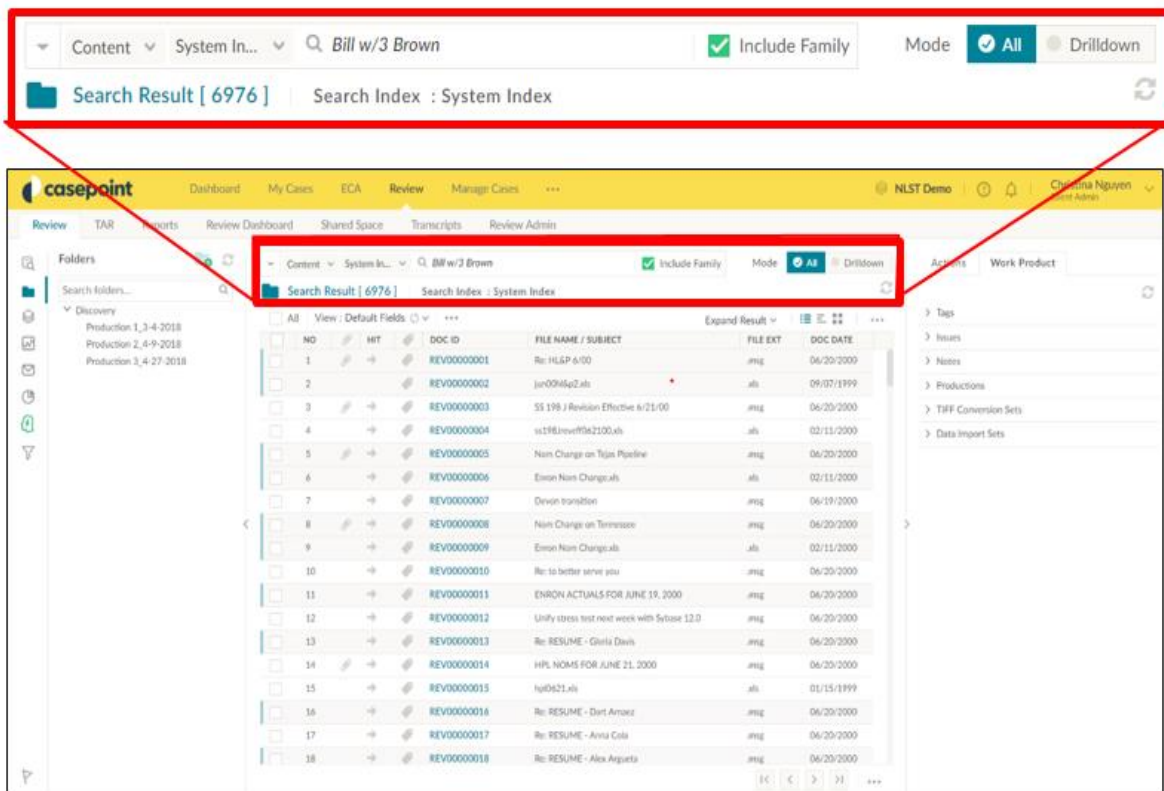
Here is an illustration of Casepoint’s home screen. This is your where you can access all or specific productions and conduct simple single word searches or complex multiple term queries.



To the right are where the Tags and Issues live. We've already created the ones we want and run searches for each one.



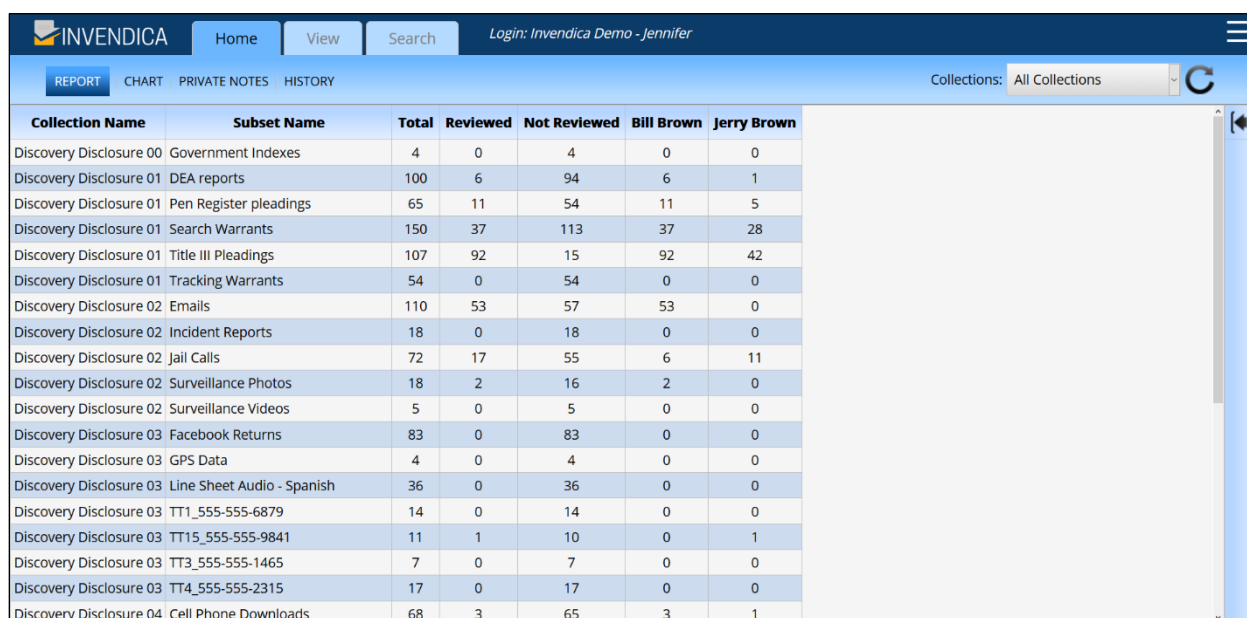
How did we search? Here's a sample search performed for Bill Brown.



The search will need to be repeated with other forms of the name like “William” or using nicknames, with all the results going into the same Tag.

Using Invendica to Search Audio and Video

Invendica is a database often used for cases with a lot of audio and video. Although it can’t search spoken words or identify faces, it can stream recordings rather than make you download files to your computer. It can also ingest formulas created by other technology companies to improve search capabilities. What makes it work is having terms that can be searched, like metadata or, for wiretaps, searchable linesheets. It has a simple home screen that displays all the productions and folders like a spreadsheet.



The screenshot shows the Invendica web application interface. At the top, there is a navigation bar with 'Home', 'View', and 'Search' tabs. Below this is a sub-navigation bar with 'REPORT', 'CHART', 'PRIVATE NOTES', and 'HISTORY' tabs. The 'REPORT' tab is active. On the right side of the sub-navigation bar, there is a 'Collections:' dropdown menu set to 'All Collections' and a circular refresh icon. The main content area displays a table with the following columns: 'Collection Name', 'Subset Name', 'Total', 'Reviewed', 'Not Reviewed', 'Bill Brown', and 'Jerry Brown'. The table lists various discovery disclosures and their associated data.

Collection Name	Subset Name	Total	Reviewed	Not Reviewed	Bill Brown	Jerry Brown
Discovery Disclosure 00	Government Indexes	4	0	4	0	0
Discovery Disclosure 01	DEA reports	100	6	94	6	1
Discovery Disclosure 01	Pen Register pleadings	65	11	54	11	5
Discovery Disclosure 01	Search Warrants	150	37	113	37	28
Discovery Disclosure 01	Title III Pleadings	107	92	15	92	42
Discovery Disclosure 01	Tracking Warrants	54	0	54	0	0
Discovery Disclosure 02	Emails	110	53	57	53	0
Discovery Disclosure 02	Incident Reports	18	0	18	0	0
Discovery Disclosure 02	Jail Calls	72	17	55	6	11
Discovery Disclosure 02	Surveillance Photos	18	2	16	2	0
Discovery Disclosure 02	Surveillance Videos	5	0	5	0	0
Discovery Disclosure 03	Facebook Returns	83	0	83	0	0
Discovery Disclosure 03	GPS Data	4	0	4	0	0
Discovery Disclosure 03	Line Sheet Audio - Spanish	36	0	36	0	0
Discovery Disclosure 03	TT1_555-555-6879	14	0	14	0	0
Discovery Disclosure 03	TT15_555-555-9841	11	1	10	0	1
Discovery Disclosure 03	TT3_555-555-1465	7	0	7	0	0
Discovery Disclosure 03	TT4_555-555-2315	17	0	17	0	0
Discovery Disclosure 04	Cell Phone Downloads	68	3	65	3	1

Invendica subject matter categories are just that - “categories” and not folders or tags. Although you can search within a category and create more categories, they do not display as folders with subfolders. They display linearly on the same category level.

Collection Name	Subset Name	Total	Reviewed	Not Reviewed	Bill Brown	Jerry Brown
Discovery Disclosure 00	Government Indexes	4	0	4	0	0
Discovery Disclosure 01	DEA reports	100	6	94	6	1
Discovery Disclosure 01	Pen Register pleadings	65	11	54	11	5
Discovery Disclosure 01	Search Warrants	150	37	113	37	28
Discovery Disclosure 01	Title III Pleadings	107	92	15	92	42
Discovery Disclosure 01	Tracking Warrants	54	0	54	0	0
Discovery Disclosure 02	Emails	110	53	57	53	0
Discovery Disclosure 02	Incident Reports	18	0	18	0	0
Discovery Disclosure 02	Jail Calls	72	17	55	6	11
Discovery Disclosure 02	Surveillance Photos	18	2	16	2	0
Discovery Disclosure 02	Surveillance Videos	5	0	5	0	0
Discovery Disclosure 03	Facebook Returns	83	0	83	0	0
Discovery Disclosure 03	GPS Data	4	0	4	0	0
Discovery Disclosure 03	Line Sheet Audio - Spanish	36	0	36	0	0
Discovery Disclosure 03	TT1_555-555-6879	14	0	14	0	0
Discovery Disclosure 03	TT15_555-555-9841	11	1	10	0	1
Discovery Disclosure 03	TT3_555-555-1465	7	0	7	0	0
Discovery Disclosure 03	TT4_555-555-2315	17	0	17	0	0
Discovery Disclosure 04	Cell Phone Downloads	68	3	65	3	1

Using Invendica for Wiretap Search and Review

Gang cases often have tens of thousands of wiretap interceptions. Here is an example of locating and saving key interceptions using Invendica.

Let's assume your client's phone was the target of a wiretap intercept and you've identified the number as 555-555-6879. You can search for your client's number in the Target Number field. The search is done using only the last four digits of the number. Agents will often format phone numbers differently, with or without hyphens, parentheses, or spaces. Searching only the last four digits will capture as many results as possible. This search brings back 13 results. Create a category for your results. Let's call it "Wires 6879".

Collection	Subset	File Name	Subject	Notes	Extension	Line Sheet	Audio	Case	Target Number	Session
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheet 1			pdf			RY-09-0045	555-555-6879	1
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 10						RY-09-0045	555-555-6879	10
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 11						RY-09-0045	555-555-6879	11
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 12						RY-09-0045	555-555-6879	12
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 13						RY-09-0045	555-555-6879	13
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 2						RY-09-0045	555-555-6879	2
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 3						RY-09-0045	555-555-6879	3
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 4						RY-09-0045	555-555-6879	4
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 5						RY-09-0045	555-555-6879	5
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 6						RY-09-0045	555-555-6879	6
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 7						RY-09-0045	555-555-6879	7
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 8						RY-09-0045	555-555-6879	8
Discovery Disclosure 03	TT1_555-555-6879	TT1_555-555-6879_Line Sheets 9						RY-09-0045	555-555-6879	9

The Dialed Digits field contains the phone number in contact with the Target Number. By adding a search for the last four digits “3602” in the Dialed Digits field, the 13 results from the previous search were reduced to 1 result.

The screenshot shows the INVENDICA search interface. On the left, the 'Custom query search' section has 'Target Number' set to 'Like 6879' and 'Dialed Digits' set to 'Like 3602'. The main table shows a single result with 'Dialed Digits' as '555-555-3602'. A callout box labeled 'Dialed Digits' points to this value. The table headers include Number, Date, Start Time, Duration, Direction, Type, Dialed Digits, Classification, Minimized, Language, and Comments.

There may be wiretap interceptions where your client’s number was not identified on the linesheet, but your client was a participant or mentioned during the call. You can search for names and nicknames in both the Participants and Synopsis fields to identify as many wiretap interceptions relating to your client as possible. The example below shows a search for “Martinez” in either the Participants or Synopsis field.

The screenshot shows the INVENDICA search interface with 'Participants' and 'Synopsis' set to 'Like Martinez'. The main table displays 18 results. A callout box labeled 'Participants' points to the first row of the results table. The table headers include Participants, Monitored, Synopsis, Doc Type, Issue, and Fa.

Participants	Monitored	Synopsis	Doc Type	Issue	Fa
ENRIQUE MARTINEZ; DAVID RODRIGUEZ	ralbisur	IN 2325; SUBS: ENRIQUE MARTINEZ TO DAVID RODRIGUEZ			
JOSE LOPEZ; ANNA MARTINEZ	ralbisur	OUT 316010158382325; SUBS:user: ANNA MARTINEZ; JO...			
ENRIQUE MARTINEZ; IRENE IVORSON	ralbisur	IN 2325; SUBS: ENRIQUE MARTINEZ TO IRENE IVORSON			
ANNA MARTINEZ; JOHN JONES	ralbisur	IN 2325; SUBS: ANNA MARTINEZ TO JOHN JONES			
ENRIQUE MARTINEZ; ANNA MARTINEZ	ralbisur	OUT 316010158382325; SUBS: ENRIQUE MARTINEZ to...			
FRED MARKEN; ANNA MARTINEZ	ralbisur	OUT 2325; SUBS: FRED MARKEN TO ANNA MARTINEZ			
JOHN JONES; ANNA MARTINEZ	ralbisur	OUT 2325; SUBS: JOHN JONES TO ANNA MARTINEZ			
ENRIQUE MARTINEZ; DAVID RODRIGUEZ	ralbisur	IN 2325; SUBS: ENRIQUE MARTINEZ TO DAVID RODRIGU...			
JOSE LOPEZ; ANNA MARTINEZ	ralbisur	OUT 316010158382325; SUBS:user: ANNA MARTINEZ; JO...			
ENRIQUE MARTINEZ; IRENE IVORSON	ralbisur	IN 2325; SUBS: ENRIQUE MARTINEZ TO IRENE IVORSON			
ANNA MARTINEZ; JOHN JONES	ralbisur	IN 2325; SUBS: ANNA MARTINEZ TO JOHN JONES			

Now you’ve culled down tens of thousands of wiretaps to a reviewable number.

Collection Name	Subset Name	Total	Reviewed	Not Reviewed	Client Review	Rank 1	Rank 2	Rank	Wires 6879
Discovery Disclosure 03	Facebook Returns	83	0	83	0	0	0	0	
Discovery Disclosure 03	GPS Data	4	0	4	0	0	0	0	0
Discovery Disclosure 03	TT1_555-555-6879	14	13	1	1	0	0	0	13
Discovery Disclosure 03	TT15_555-555-9841	11	5	6	0	0	0	0	5
Discovery Disclosure 03	TT3_555-555-1465	7	0	7	0	0	0	0	0
Discovery Disclosure 03	TT4_555-555-2315	17	7	10	0	0	0	0	7
Discovery Disclosure 04	Cell Phone Downloads	68	0	68	0	0	0	0	0
Discovery Disclosure 04	Lab Reports	42	0	42	0	0	0	0	0
Discovery Disclosure 04	N-1246	3	0	3	0	0	0	0	0
Discovery Disclosure 04	N-395	5	0	5	0	0	0	0	0
Discovery Disclosure 04	N-5300	2	0	2	0	0	0	0	0
Discovery Disclosure 04	Supplemental Discovery	14	0	14	0	0	0	0	0
Discovery Disclosure 05	181217-Flynn-302-1	1	0	1	0	0	0	0	0
Discovery Disclosure 06	Supporting Documents	58	49	9	0	0	1	0	0
Discovery Disclosure 06	Viktor Yanukovich estate docs	28	0	28	0	0	0	0	0
Discovery Disclosure 07	Wire apps and orders - Carter Page	8	0	8	0	0	0	0	0
Discovery Disclosure 08	Exhibits	29	11	18	0	0	0	0	
		1097	95	1002	1	0	1	0	25

Click on the discovery item and the linesheet will display. If there is an associated audio recording, it'll be paired with the linesheet. You won't waste time looking for the recording for a linesheet. They're paired together on the database.

Documents - 25

ID	Name	PDF
4359	TT1_555-555-6879_Line Sheet 1	PDF
4360	TT1_555-555-6879_Line Sheets 10	PDF
4361	TT1_555-555-6879_Line Sheets 11	PDF
4362	TT1_555-555-6879_Line Sheets 12	PDF
4363	TT1_555-555-6879_Line Sheets 13	PDF

Category

- ☒ Client Review
- ☐ Rank 1
- ☐ Rank 2
- ☐ Rank 3
- ☒ Wires 6879

Information

File Name : TT1_555-555-6879_Line Sheet 1

Subject :

Notes :

Extension : pdf

Line Sheet :

Audio :

TT1_555-555-6879_Line Sheet 1 Audio

Case: RY-09-0045 Line: 555-555-6879 Session Number: 001

Date: 3/18/08 Start Time: 16:38:36

Duration: 00:01:02 Direction: Incoming

Type: Voice Dialed Digits: 555-555-6879

Classification: Pertinent Minimized: No

Language: Spanish

Comments: IRENE IVORSON TOOK 200 FROM ELIZABETH'S MONEY ON ACCIDENT 2325

Participants: IRENE IVORSON; MARCO

Monitored By: ralbisur

Synopsis: IN 316010158382325; SUBS: IRENE IVORSON to MARCO

Download Native

Sound confusing? It can be when looking at a demonstration site but once you're on a database with your discovery, everything seems to change. You become more focused. When you run your first search and it comes back with what you envisioned, the confidence you'll gain will have you building off that search over and over again. It could be a matter of changing the phone numbers or names or adding a date referenced in an overt act of the indictment.

Using Invendica to Review Social Media Private Messages

Our more recent challenge is finding an easier way to review private social media messages. For those of you who are not Facebook or Instagram users, both services provide private messaging options for account holders. Messages display in a back and forth manner like a text message. They are not public to anyone else but the people messaging. The manner in which Facebook produces the materials is voluminous and not in an easily recognizable form. If your client is a long time social media user, expect tens of thousands of pages of materials.

The subpoenaed materials are not produced as printouts or screenshots of the familiar message format. Instead, the materials are provided as data where every communication, photograph, video, and account holder has a numeric identifier.

Photographs, audio, and videos are often embedded in the messages as a form of communication. In some of the subpoenaed Facebook productions, photographs do not display as part of the message, and audio and video cannot be played from these productions. To review these, you must search through a separate folder to locate the corresponding files. Piecing together a complete communication is time consuming. Piecing hundreds or thousands of communications together is unrealistic. You would spend too many hours trying. Here is an example, using demonstration data, of how Facebook private messages are produced.

Thread 555552366841710300949128191872200313231

Current 2018-11-10 09:00 SEA

Participants Sally Smith (5553595)

Brian Brake (5555584554)

Sent 2018-11-10 09:02 AM SEA

Deleted

Body Hey

Author Brian Brake (5555584554)

Sent 2018-11-10 09:10 AM SEA

Deleted

Body Hey, wyd tonight??

Author Sally Smith (5553595)

Sent 2018-11-10 09:11 AM SEA

Deleted

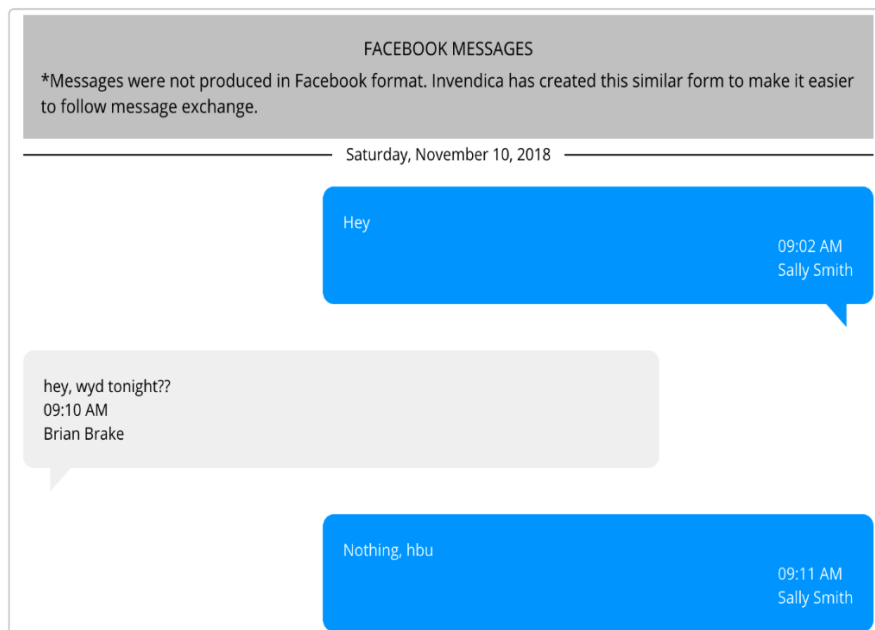
Body Nothing, hbu

Instagram private messages are produced in a similar manner.

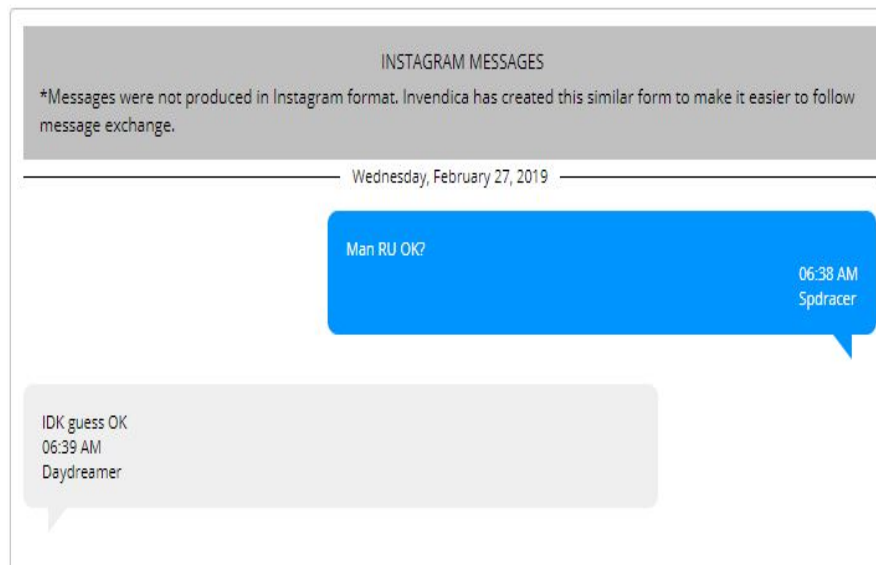
Thread-id 555552366841710300949128191872200313231
Text Man RU OK?
Author Spdracer (5553595)
Recipients Spdracer (5553595)
Daydreamer (5555584554)
Id 55555734804646900355968662922854
Time 2019-02-27 06:39:33 SEA
Item-type text
Thread-id 555552366841710300949128191872200313231
Text IDK guess OK
Author Daydreamer (5555584554)
Recipients Spdracer (5553595)
Daydreamer (5555584554)
Id 55555734804646900355968662922854
Time 2019-02-27 06:38:58 SEA
Item-type text

We worked with a technology company, Slate Technology for over a year to create a program that will extract private messages from the subpoena materials and convert them to a usable format. Invendica then ingests Slate Technology's work into its database and creates an easier-to-review format that is similar to the familiar Facebook or Instagram message form. Each file is an individual conversation among Facebook or Instagram users. The messages identify date, time, the message sender, and the communication content. Photos within the messages will appear. Audio and video will appear if the original format of the Facebook or Instagram record was an HTML file. Below is a sample of Slate Technology and Invendica's work using fictitious materials.

This is an example of a Facebook private message.



Instagram direct messages display in the same manner.



We have Invendica host the full Facebook and Instagram records on its database and separate collections of just the Facebook and Instagram private messages. Counsel can go straight to the messages and search by date, time, participants, and content to find key private Facebook or Instagram exchanges.

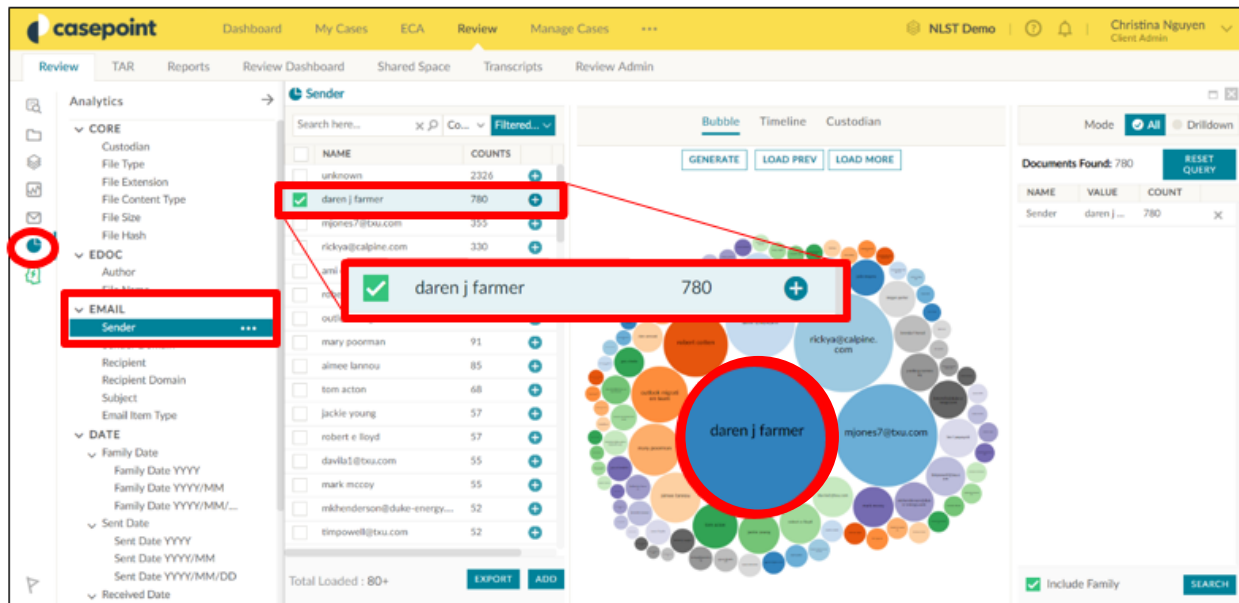
Here's a tip: The materials are typically produced in HTML and PDF format or HTML and the Government converts it to a PDF. Ask for both. The HTML version breaks the account down into sections with the messages separated. The PDF version lets you use the review features of Adobe. As you get used to the HTML version, you'll like it more.

Using Casepoint for Email Search and Review

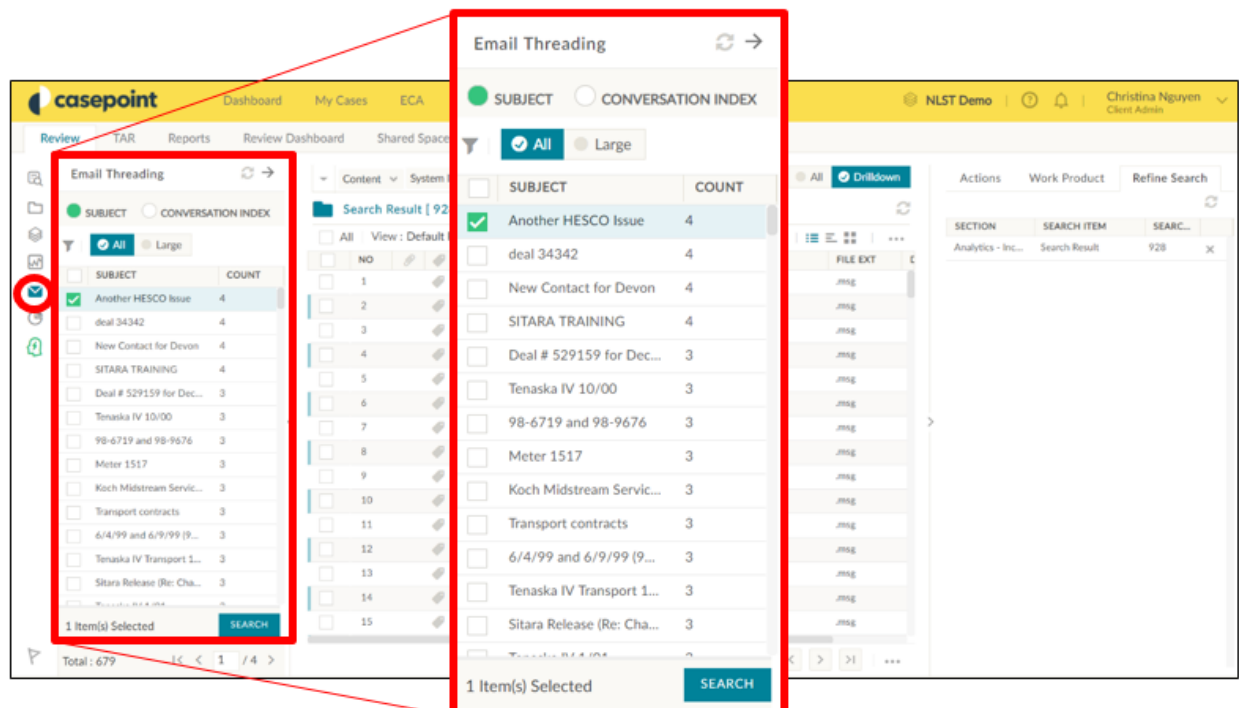
Drug cases can be huge, but the real monster discovery cases are the fraud prosecutions. Fraud cases typically involve voluminous productions of documents, search warrant pleadings, investigative reports, financial records, emails, and forensic computer drive images from computers, laptops, tablets, and cellphones. For drug cases, the wiretaps are often the most critical part of the discovery production. For fraud cases, more times than not, it's the emails. Here is an example of how a database like Casepoint can locate emails sent by a particular sender then identify the most inclusive emails in a chain.

Casepoint's analytics allows you to sort emails by various metadata fields, like "To, From, Subject, Date", who received the email, and whether it was forwarded or modified. The example below reflects emails being sorted by sender. You'll see a list of names and email addresses in the sender list. The size of the colored bubble graphs corresponds to the "counts" or amount of emails for each sender. The largest bubble is for "daren j farmer" with 780 emails where he is identified as the sender. You can add

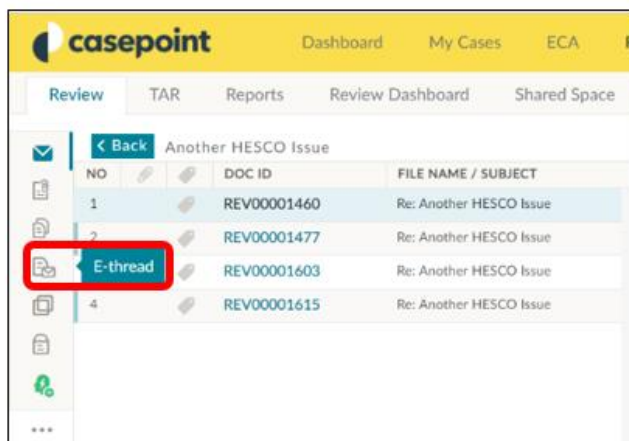
these emails to your search query by clicking the plus symbol next to the name on the vertical list. You can also double click the bubble with the name.



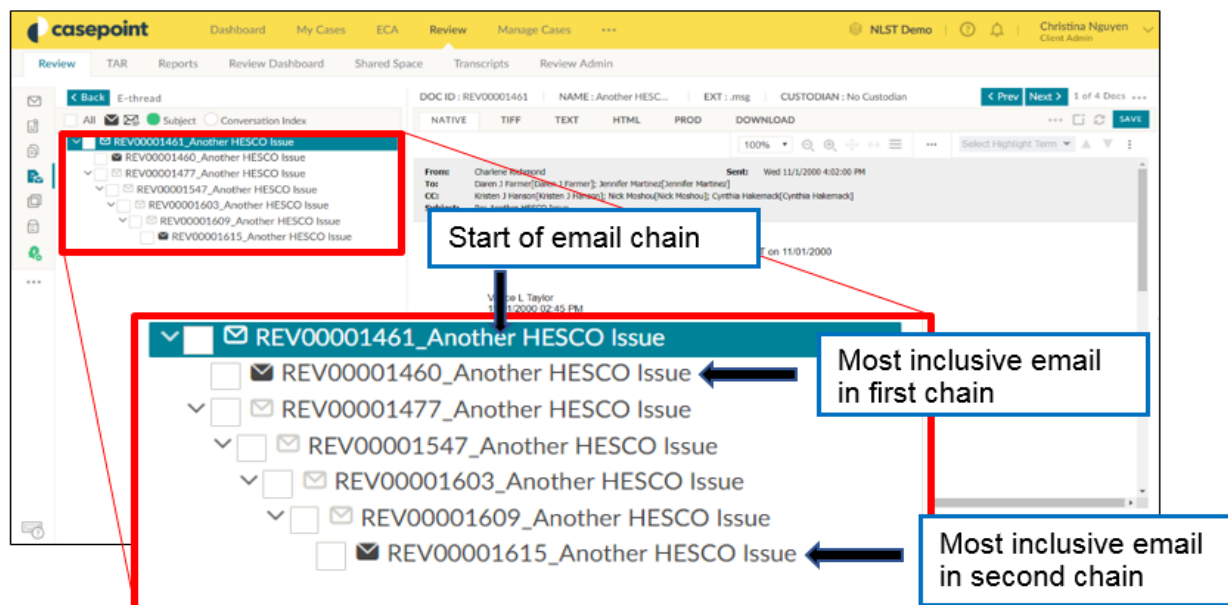
The email threading feature allows you identify emails either with the same subject line or are part of the same conversation. Below, emails with the subject line “Another HESCO Issue” have been selected.



After isolating these emails, a new icon menu bar specific to email threading will appear on the left-hand side of the screen. Highlighted below is the icon for Casepoint's E-thread function.



The E-thread function organizes the emails into a chain showing the starting email, response emails, and the most inclusive emails. The black mail icons indicate these are the most inclusive emails in their respective chains. Instead of trying to piece together fragments of an email conversation, you can go right to the most inclusive email – a huge time saver!

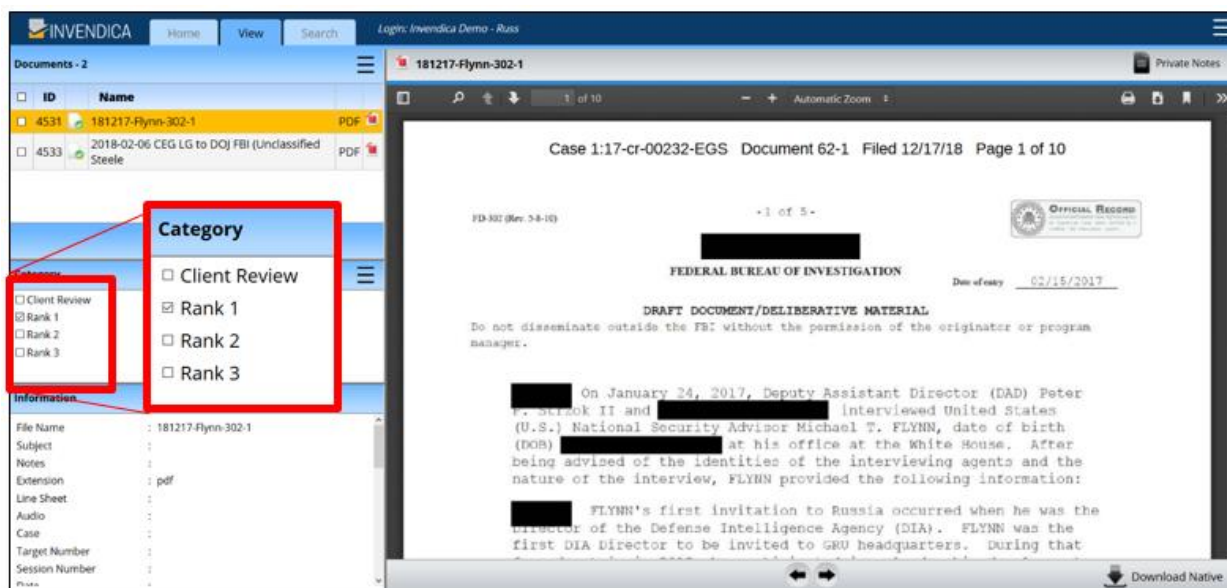


Invendica: Review your Results and Rank by Importance

At one point, after searching, sampling, revising your search, and sampling again, you'll get to where it is time to review your search results. Not all of what you found will be pertinent to your case. This is a good place to add a ranking system to your review process.

If you've reduced a single subject collection from 5,000 to 100 discovery items and you're reviewing them one by one, then for every item you find to be important to your case you could categorize it Rank 1. For items you know are not important, you could categorize them as Rank 3. For items you're not sure of, you could categorize them as Rank 2. Later, when you've completed more of your discovery review, you can go back to everything you've ranked as a "2" to reclassify them as Rank 1 or Rank 3. Eventually, you'll be working with all discovery items you have categorized as Rank 1.

Here's an illustration of how it would look on Invendica. It's a little cumbersome, but it works. As you review each discovery item, you can check the applicable box on the left.



Your totals appear on the Homepage.

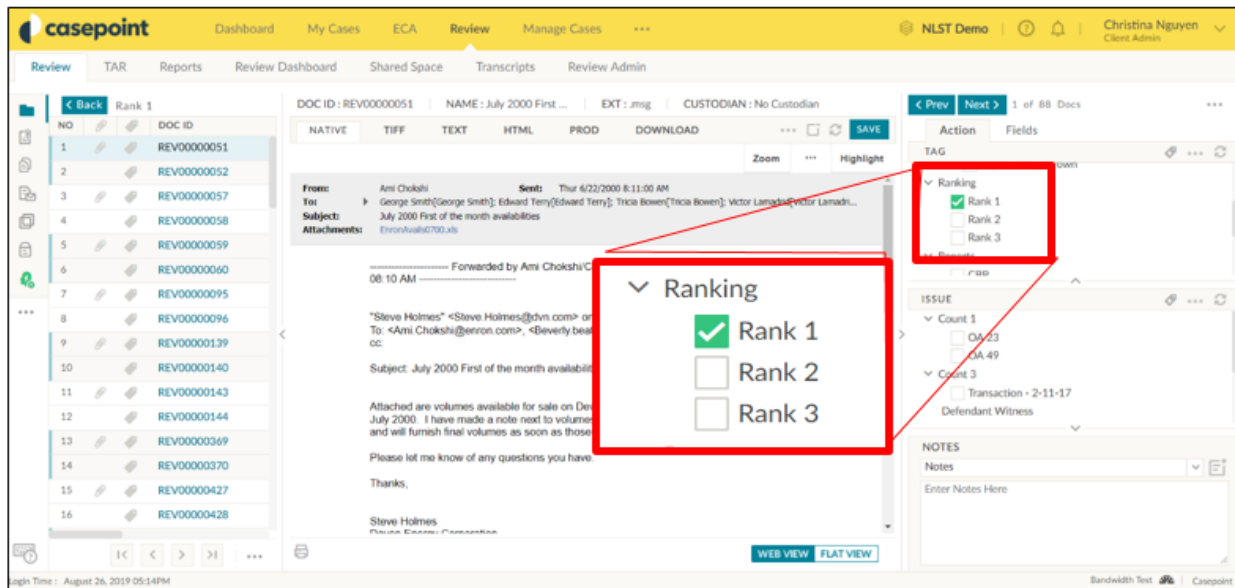
Collection Name	Subset Name	Total	Reviewed	Not Reviewed	Rank 1	Rank 2	Rank 3
Discovery Disclosure 03	Facebook Returns	83	3	80	0	3	0
Discovery Disclosure 03	GPS Data	4	0	4	0	0	0
Discovery Disclosure 03	TT1_555-555-6879	14	13	1	0	0	0
Discovery Disclosure 03	TT15_555-555-9841	11	5	6	0	0	0
Discovery Disclosure 03	TT3_555-555-1465	7	0	7	0	0	0
Discovery Disclosure 03	TT4_555-555-2315	17	7	10	0	0	0
Discovery Disclosure 04	Cell Phone Downloads	68	39	29	0	0	39
Discovery Disclosure 04	Lab Reports	42	0	42	0	0	0
Discovery Disclosure 04	N-1246	3	0	3	0	0	0
Discovery Disclosure 04	N-395	5	0	5	0	0	0
Discovery Disclosure 04	N-5300	2	0	2	0	0	0
Discovery Disclosure 04	Supplemental Discovery	14	0	14	0	0	0
Discovery Disclosure 05	181217-Flynn-302-1	1	0	1	0	0	0
Discovery Disclosure 06	Supporting Documents	58	1	57	0	1	0
Discovery Disclosure 06	Viktor Yanukovich estate docs	28	0	28	0	0	0
Discovery Disclosure 07	Wire apps and orders - Carter Page	8	0	8	0	0	0
Discovery Disclosure 08	Exhibits	29	0	29	0	0	0
		1097	116	981	48	4	39

From here, you can search all the “Rank 1” documents that appear in your “Bill Brown” then save them as “Bill Brown 1”.

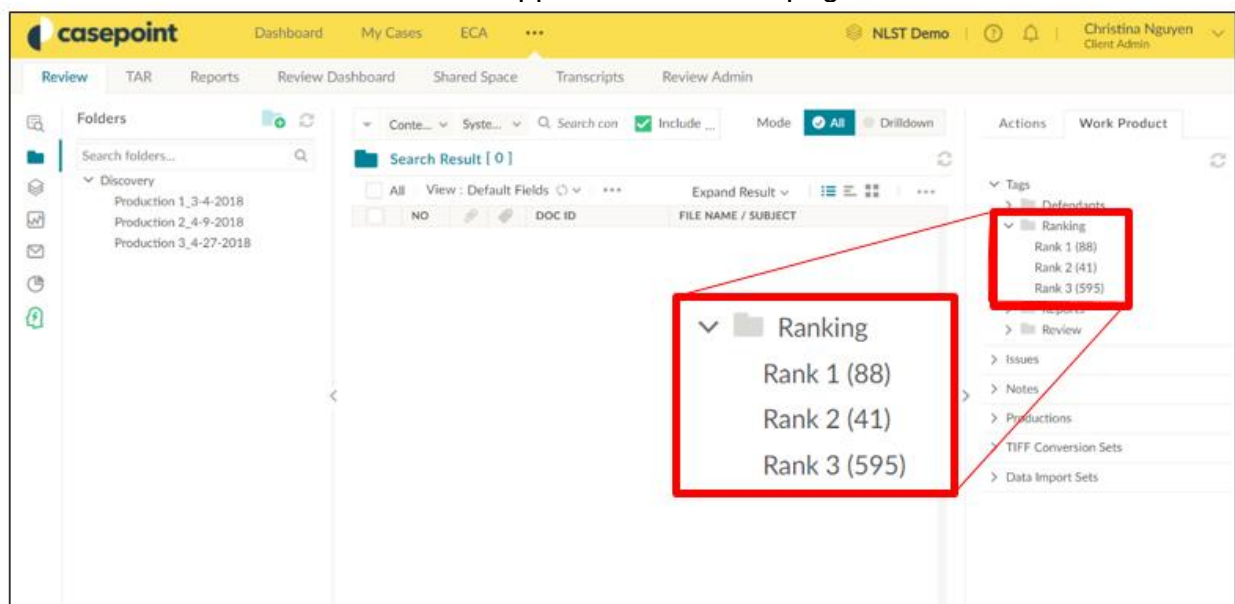
Casepoint: Review your Results and Rank by Importance

For more document-intensive cases, Casepoint allows you to create subfolders where you can rank documents.

You can check the box for the ranking that applies and save that categorization.



As with Invendica, the totals appear on the homepage.



Once you're down to all "Rank 1" discovery items, you can do the same analysis you use for all your cases to determine how these documents can work for your defense. But instead of a sticky note on a stack of documents, you can categorize all the discovery items on the database for things like impeaching a particular witness, addressing a money laundering charge, or any other topic of importance. When it comes time for trial, you can download them all to a thumb drive.

Conclusion

These are just a few examples of how you can use a database to handle cases where there is more discovery than you can ever possibly review. A database is a powerful tool. Invendica provides the ability to search across all the wiretap interceptions to pull up both the audio recordings and the linesheets. It can also associate transcripts to the linesheets and audio recordings. Casepoint provides powerful analytics including email threading and, for the more adventurous, Technology Assisted Review (predictive coding).

For any of these platforms and regardless of what you may be looking for, the same strategy we started off in this memo applies.

- Identify broad subjects you want to research and the search terms to use (like all items that mention your client)
- Search, sample, refine your search, sample again
- Save your results as general subject matter categories
- Search within your general subject matter categories for more specific items using the search, sample, search method
- Then review your results and rank the results based on importance (Rank 1, Rank 2, Rank 3)
- Review all your Rank 1 items as you would for any case, and identify which items go to your defense, particular witnesses, overt acts, or counts in an indictment

In contrast, manually looking at every file produced, one item after another, would be mind-numbing work. Spare your mental energy. Cut to the chase by using technology to reduce huge discovery productions down to manageable collections, then apply the same review strategy you've always used. Instead of trying to stay alert for 50,000 discovery items, you may only need to be the most attentive for 500 files – a huge difference with a higher likelihood of not spacing out and missing important discovery items.

Need more help? Contact Defender Services' National Litigation Support Team or any of the five Defender Services Coordinating Discovery Attorneys around the country.