

🖨 [Click to print](#) or Select 'Print' in your browser menu to print this document.

Page printed from: <https://www.law.com/thelegalintelligencer/2019/06/12/processing-and-review-what-comes-after-ilc/>

Processing and Review: What Comes After ILC?

In this, the third installment of our e-discovery series, we're going to bring things up to speed and jump back in line with what happens after electronically stored information (ESI) has been “preserved, identified, located and captured.”

By **Patrick Kennedy** | June 12, 2019



Patrick

When last we left our heroes, they were talking about preservation

Kennedy, Chamberlain Hrdlicka

(<https://www.law.com/thelegalintelligencer/2019/03/15/preservation-you-cant-produce-what-you-dont-have/>). That's the step right before "identify, locate, capture" (ILC (<https://www.law.com/thelegalintelligencer/2019/01/29/esi-in-2019-the-e-stands-for-so-much-more-than-just-email/>)), meaning that the second article served as more of a "prequel." In this, the third installment of our e-discovery series, we're going to bring things up to speed and jump back in line with what happens after electronically stored information (ESI) has been "preserved, identified, located and captured." That is the process of ... processing! Once that's done, it's time to review what we have. In these two steps, the aim is to take all of the ESI from the prior steps and put it in a format that can be quickly and logically reviewed by attorneys on both sides.

The capture part of the procedure is designed to cast a wide net because, as we stated, you can't produce what you don't have

(<https://www.law.com/thelegalintelligencer/2019/03/15/preservation-you-cant-produce-what-you-dont-have/>). The capturing or collection process is more efficient when gathering as many files as you can the first time you encounter them—this allows just a one-time visit with the end user. But just like your closet or your garage collects a lot of stuff that you don't use or need anymore, the collection phase of e-discovery also yields files that are unusable or unneeded. Think, for example, about executable files—the files that are necessary to run software on your computer. If you want to open a Word document, you click the icon on your task bar that runs a file called word.exe. That's a necessary file for sure, but not something that is going to be of any use to an attorney or to a client. This is exactly the kind of file that would need to be culled out during the processing phase of the electronic discovery reference model (EDRM). Once you gather the ESI from the sources you identify, the processing phase is the next step so that attorneys end up reviewing only the files that could potentially matter to the case in the most efficient way possible.

The issue with ESI is that it comes in from all different sources and applications (think back to all the applications you use, and where that data is stored). In addition to culling out unneeded files like executables, figuring out how to look at the rest of the

data and make sense of it is what e-discovery processing software does: it takes unstructured data and turns it into structured data. What does that mean, exactly? Let's dive in.

Unstructured data simply refers to data that either does not have a pre-defined data model or is not organized in a pre-defined manner. Unstructured information is typically text-heavy, but may contain data such as dates, numbers and facts. A Microsoft Word document, a PDF file, a PowerPoint presentation, all of these would be considered unstructured data. Now, think about what parts of the files we really need—the text-based information. What processing does is pull text out of unstructured data sources and put it into a database format. Why? Because database formats are structured. They consist of tables with fields that can be sorted, filtered, grouped and categorized. Moving to a database format allows us to take a large volume of ESI and put it in neat rows and columns so that we can flag each file as relevant, confidential or a host of other categories that help determine what needs to get produced to the other side.

To put more structure (see what I did there?) to how that happens, I've broken it down into steps that are taken during processing:

- Import the ESI into the processing tool.
- Strip all text and metadata from each file (metadata includes information about the file, such as date created, date modified, the size of the file, etc.).
- Place the data from step two into rows in a database, one for each electronic file.
- Give each row a unique ID (document ID, serial number, etc.).
- Give each row a "pointer" to the original file that the text and metadata was stripped from.
- Create an export load file and export the processed data from the processing engine and save it to a safe location with a short user access list.
- Load data by volume, into the review tool, so that attorneys can review each file.

When do you move to a database that allows structured review? That threshold will vary from case to case, but a general rule that we use at Chamberlain, Hrdlicka, White, Williams & Aughtry is when a collection reaches over 1,000 files, it's time to use a

processing and review tool.

No discussion about e-discovery processing would be complete without a word about errors. When you bring data into the processing tool, and potentially again when you create a load file for the review tool, there are often errors that occur. Processing tools call these “exceptions,” though that does not mean exceptional—not in the way you would think. The issue is, with the many different file formats in use today, not all of them will go quietly into that dark night of your processing tool. Some files will be troublemakers, and there’s nothing you can do about it. You need to make sure that your vendor or in-house staff understands what you want done with those files. Keep in mind that the “smoking gun” may in fact be in a set of ESI that refused to cooperate. Handling exceptions is a one-by-one process in and of itself, and can add to the time, and therefore costs, of processing.

Once the processing phase is complete, and exceptions have been handled, what happens next? This is the review phase, the most important step in the process, because this is when to make the decision about whether a file should be produced to the opposing side. If you envision the entire EDRM as a funnel, the amount of ESI is reduced in each phase. By the time you get to the review phase, near the tip of the funnel, attorneys are only looking at files that they should actually review.

Prior to attorneys reviewing each file in the database via a review tool, both sides should agree on a set of general categories into which all files can be placed. When it comes time to produce to the other side, it will then be a matter of checking off which categories of data need to be sent over and which don’t. To reiterate from the above section on structured versus unstructured data, this is why we move to a database: running a filter on all rows of data that are marked “responsive” can be done very quickly.

There will always be at least two tiers of categories that all ESI falls into when being reviewed:

- Tier 1—Mutually exclusive

- Responsive
- Not Responsive

- Tier 2—Case-specific multipick
 - Users can pick more than one category from this list, as more than one may apply to a single individual file/document.

Sidebar: Many review platforms call this process of categorizing data “tagging”—e-discovery consultants set up “tags” for documents. For clarity, this article refers to them as “categories.”

What follows is an example of what review categories might look like for a case involving construction (<http://www.constructionexec.com/article/limit-litigation-expenses-by-streamlining-storage-of-electronic-informatio>). Tier 1 categories are repeated here as a reminder that those are required categories for every file in all document reviews:

- Tier 1
 - Responsive
 - Not Responsive

- Tier 2
- Contracts
- Emails between Custodian A and B
- Emails between Custodians A and C
- Drawings and design documents

Reviewing ESI is meant to reduce the amount of data passed to and from each side down to only what is responsive or relevant to the case. It also helps ensure that ESI gathered during the collection phase that should not be exposed to opposing counsel is, in fact, not. Setting up categories and marking or “tagging” documents to place them into those categories aids in this process tremendously. By processing data first—both

to remove unwanted and unneeded file types and to move into a structured database format—and then reviewing and categorizing it, we can get a clear handle on what needs to be produced to the other side.

Exactly how that production step works is the subject of the last article in the series. You'll want to tune in to see how the story ends!

Patrick Kennedy *is the director of e-discovery for Chamberlain, Hrdlicka, White, Williams & Aughtry, a multidiscipline law firm with offices in Philadelphia, Atlanta, Houston and San Antonio. Contact him at patrick.kennedy@chamberlainlaw.com (<mailto:patrick.kennedy@chamberlainlaw.com>).*

Copyright 2019. ALM Media Properties, LLC. All rights reserved.