



**Digital Forensics and eDiscovery**

# **A Guide to End-to-End Digital Investigations**

## Introduction

# Understanding Digital Investigations



As a blanket statement, it is safe to say that almost any organization active in the current legal and technological era will find itself performing a digital investigation. To be sure, most (if not all) information stored by organizations nowadays is digital, or has a digital copy stored somewhere. As a result, any organization that stores any type of information will have to perform some type of digital investigation, or should at the very least have the infrastructure and policies in place to perform one.

The depth and range of this infrastructure depends on a range of factors including (but not limited to) the types of data stored, the amount of data stored and the legal/regulatory environment the organization is active in. As the depth and range of the infrastructure increases, as does the need for tools and manpower to execute them.

As data volumes [continue to grow](#), so does the complexity associated with investigating that data. Whether it's for an internal investigation or for compliance to an external regulatory request, the days of CTRL-F searches are well and truly behind us.

With this playbook, our purpose is to present the best practices and takeaways from our experiences with digital investigations, to help organizations create a predictable, repeatable and accountable investigative process.

We aim to help organizations avoid ad-hoc decision-making, and provide a clear path to conducting productive, solution-based digital investigations.



### **ZyLAB, an IPRO company**

ZyLAB believes technology can help perform fast and comprehensive digital investigations. Since Legal exposure hides everywhere, investigators need solutions to discover risk effectively.

# The Investigative Framework

The Electronic Discovery Reference Model (EDRM) is a reference process widely adopted by digital investigators, especially in the United States. The EDRM (see Fig. 1) provides the framework for the eDiscovery cycle and consists of nine phases; four dealing with on-premises identification, preservation, and collection of data (also referred to as the “left-side”) and five (referred to the “right-side”) that deal with the analysis, preparation, and presentation of data.

The framework of eDiscovery, part of litigation-related evidence gathering, is very applicable to digital investigations in general.

In addition to this framework, there are a few terms that will be used worth explaining in detail.

## Electronically Stored Information (ESI):

refers to information created, manipulated, communicated, stored or utilized in digital form (i.e., on computer equipment, servers, hard drives, personal digital assistants, smartphone devices, back-up tapes, sensors, web-based storage, etc.). ESI includes, among other things, employee email, home directories and chat logs. Next to these, sources are groups directories, payments systems, log files and other production systems.

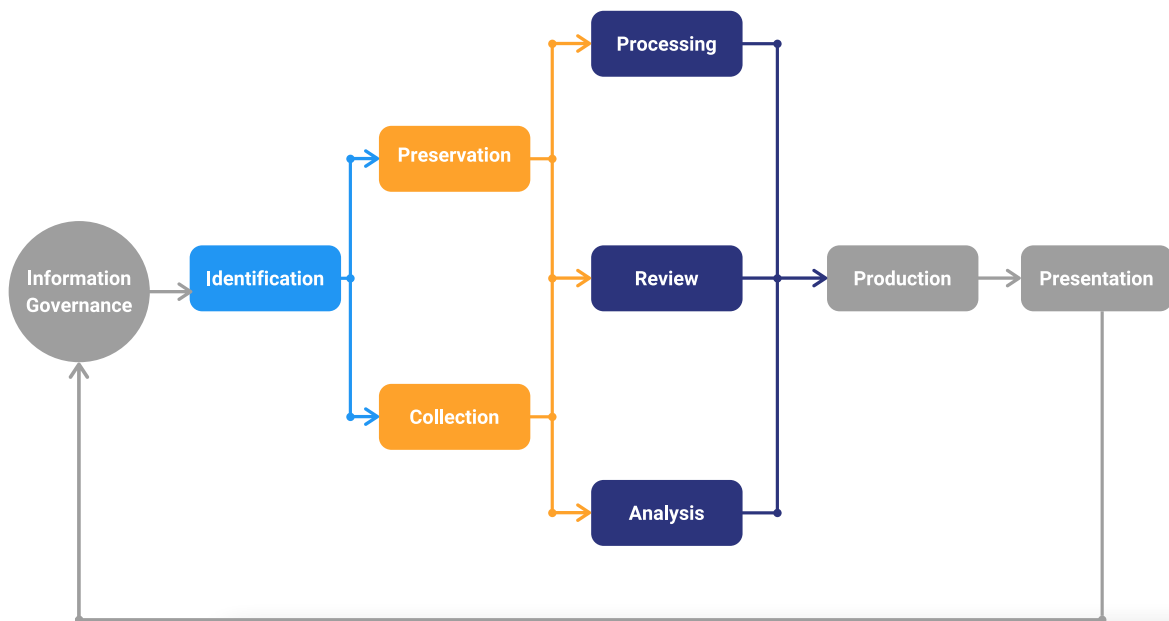


Fig. 1 - Electronic Discovery Reference Model



The difficulties of searching unstructured data are at the center of most challenges in Digital Investigations.

**Structured and unstructured data:** not all ESI is the same. Of the seven data types organizations deal with, the relevant two for investigations are structured and unstructured ESI. Structured ESI refers to data held in databases.

Usually this information is relatively easy to search as the database provides a framework (or a structure, hence the term) that organizes the information it holds. Structured data is generally easy to access and search, both by computers and people. Unstructured data is the opposite - ESI held in the aforementioned emails, drives, phones, web storage, etc. ESI of this type is both more common and harder to search. Depending on the type of investigation, unstructured ESI might be the primary or even only focus of the search. When discussing the challenges digital investigations face, most issues are centered around the challenges of searching unstructured data.

**Data Retention:** refers to the process utilized to manage documents from creation to destruction (also referred to as the record life cycle). Generally, the principle criteria that governs a record life cycle is the purpose for which the record is created and utilized. Retention programs, also known as record management programs, most often direct that records should be discarded once they are no longer of use to an organization. The scope and nature of retention programs and schedules are based on its regulatory retention obligations, operational needs, resources and risk tolerance.

**Data Preservation:** refers to the preservation of information and records specifically for purposes of legal matters, regulatory investigations and/or third-party subpoenas. Preservation ensures that information which is potentially relevant to a case remains available as it is handled, notably by exempting that data from retention policies.

# Digital Forensics, eDiscovery, or Both?

Widespread digitization, coupled with the affordability of storage has resulted in huge growth in the volume of digital information being created and stored. When it comes to digital investigations, the terms eDiscovery and digital forensics are often used interchangeably. Though the two may seem similar at first glance, there are key differences between the two. Which isn't to say they are opposites. Rather, they are complementary tools digital investigators can use to complete their objectives.

Being able to tell the difference between a forensic and eDiscovery challenge, and where the interchange between the two occurs, is key to conducting effective and comprehensive digital investigations.

## Digital Forensics

Extracting digital information from devices is a technical task, requiring appropriate tools, training and experience to be done correctly. Digital forensics becomes involved when the ESI in question is made unavailable or inaccessible: a forensic specialist can make this data available or accessible for investigation. The job of a forensic specialist is recovering deleted files, investigating manipulated or falsified information, analyzing whether and how intellectual property has left your organization, etc. In short, the digital forensic process is about identification, preservation, acquisition, processing, extraction, analysis and documentation of digital evidence. To put it in the simplest of terms: digital forensics is about finding the ESI.

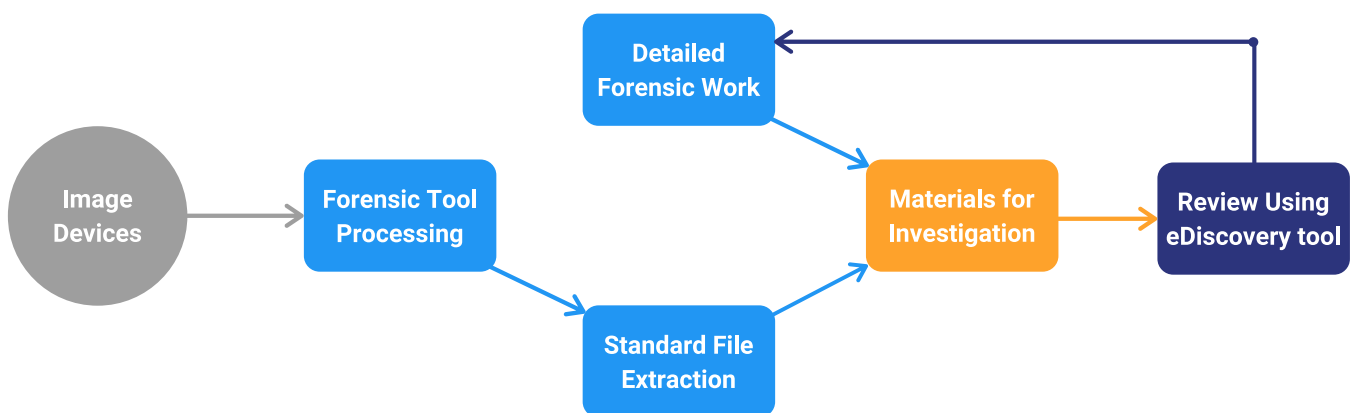


Fig. 2 - The digital forensics process

## eDiscovery

eDiscovery helps in the sorting, searching and review of available data. These tools generally aim to collect and search available and extractable information for potential evidence. As the EDRM illustrates, the 'source' for this information is [information governance](#), a system of policies and tools that governs known available ESI. Therefore, eDiscovery is best applied when there's no reason to suspect that data has been intentionally deleted, damaged or hidden. Once data is fed into the eDiscovery tool, it possesses advanced techniques to organize (impose structure) unstructured data and search it. For the purposes of data investigations, the EDRM flow can be simplified to reflect the absence of American civil procedure requirements. To again apply the simplest of terms: eDiscovery is about what's contained in the ESI.

## Combined efforts

Although used interchangeably, digital forensics and eDiscovery aren't quite as similar as they might appear at first glance. Rather, they should be considered part of the complete scope of a digital investigation. For internal investigations especially, only going off the ESI that's readily available and extractable is ill-advised: evidence may have been deliberately hidden or removed.

Again, it's worth remembering that eDiscovery is a product of the rules formulated for civil litigation in the United States, which explains why the EDRM only considers Information Governance as a source of ESI. It is, at its core, a tool for legal self-defense when involved in civil litigation. As such, the model has a few blind spots for uses outside its traditional scope. Namely, it assumes all information is available.

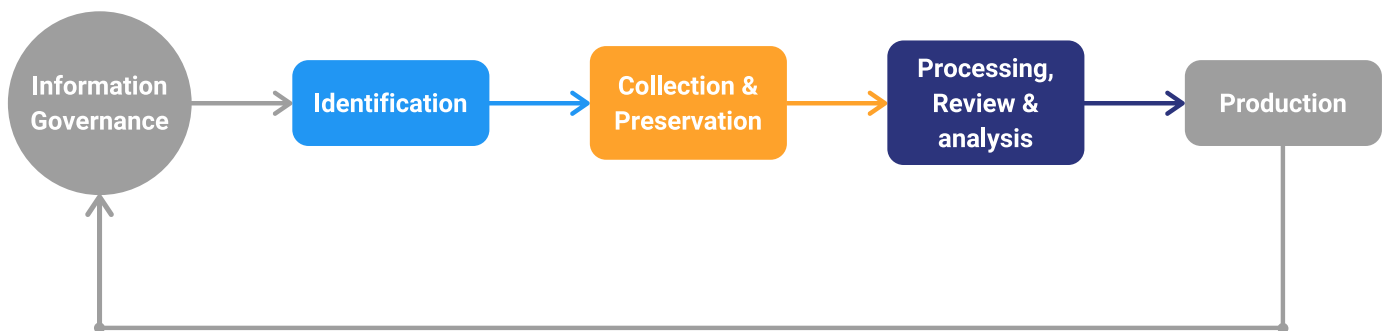


Fig. 3 - eDiscovery, simplified

Amending the simplified EDRM to include the possibility of data having to be recovered by digital forensic efforts would combine some of the terms used in each model, creating an end-to-end model where the top half represents work done in tools for digital forensics and the bottom in eDiscovery tools. This end-to-end model for investigations combines the capabilities of each into a single process.

As shown in figure 4, the combined processes would see both information governance and image devices contribute ESI to the process, with eDiscovery's data identification stage and

digital forensics' Standard File Extraction stage overlapping in a stage named 'File extraction'. Following this extraction stage, eDiscovery tools are used to collect and search the totality of the potentially relevant ESI. If at any point during the investigation information appears to be missing, more detailed forensic efforts can be made to recover this data.

By combining the tools of digital forensics and eDiscovery, digital investigators have a complete set of tools at their disposal to ensure all potential data is subject to their investigation, and can be searched thoroughly and effectively.

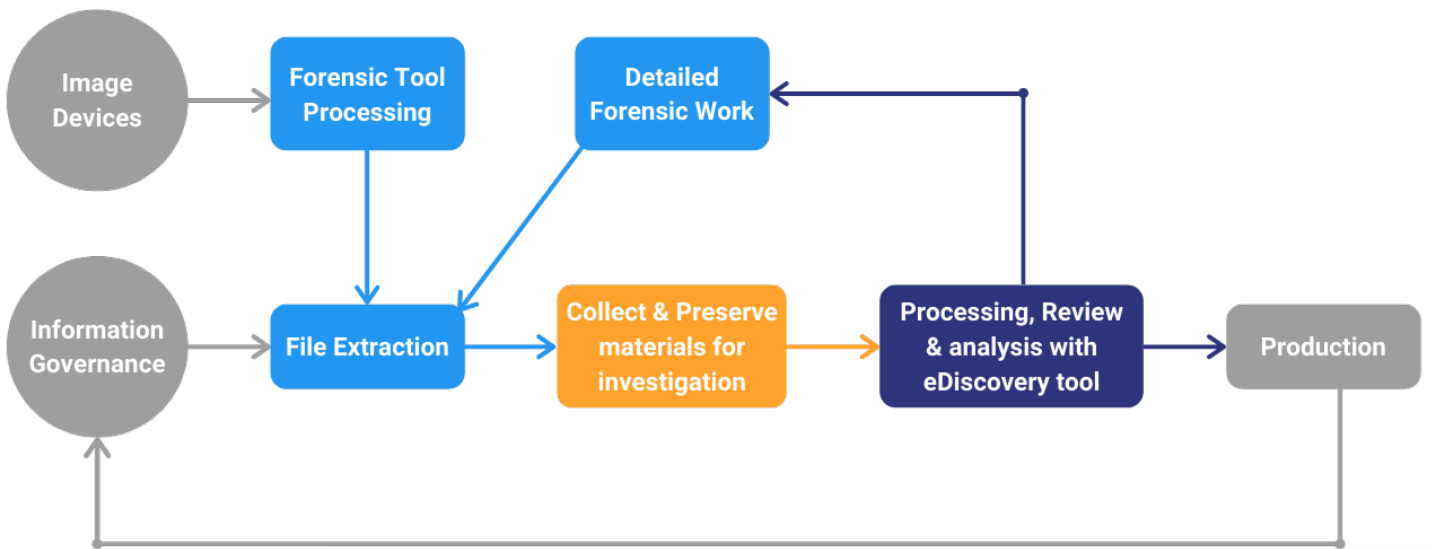


Fig. 4 - Digital forensics and eDiscovery, combined





### Chapter Three

## Finding & Identifying Data Sources

The first order of business for an investigation is to identify potential sources of relevant information. Nowadays, most (if not all) data will be ESI. Finding the exact sources, typically means speaking to key players in order to find out what type of relevant records they may or may not be holding. IT staff and, if applicable, records management personnel, should be consulted in order to establish storage locations, retention policies, data accessibility, and the availability of tools to assist in the identification process.

The process of mapping data sources and collecting data revolves around answering three major questions:

- Where can the ESI that we are looking to collect be found?
- How will we collecting potentially relevant ESI?
- Who should be involved in the process of data collection?

Once these questions have been answered, identification is complete, and the investigation can begin with said data. From a case management point of view, executing a clear and defensible investigation is paramount, regardless if it concerns an internal investigation, an audit, legal issues, or regulatory requests.

## Identifying potential data sources

Data may be stored in any number of sources including email, computers, mobile devices, databases, tape backups, and third-party sources such as cloud storage and cloud backup sites. The data maps of organizations contain a multitude of [data silos](#), a complex network of applications, tools and databases that are not necessarily connected to one another. This

custodians, questionnaires can be sent to the custodians to confirm that all potential sources of information are identified. Including users and asking them what they use is crucial, one key finding of a [2017 survey](#) was that the average worker uses 9.4 apps every day, and 48% of those applications are not provided by IT. Only going off what IT is aware of may lead to key data being missed.



48% of applications in use in organizations are not provided by the IT department.

is a challenge now, and one that continues to grow as the number of applications in use by organizations is [growing fast](#): from an average of 16 apps in 2017 to 80 in 2020. In 2021, the average company used 110 applications. Mapping out this complex environment and understanding what information goes where is key to conducting successful investigations.

The investigators should make use of appropriate legal and IT resources to determine where relevant data may be located. Data can also be located exclusively in user terminals, or be held by specific individuals. Such individuals are called custodians. If it is possible to inform and involve

That means once custodians are identified, learning where they store the data they use professionally is vital to completing the map of potential data sources for an investigation.

### Appropriate methods for data collection

Prior to collection, the trade-offs involved in different collection methods should be considered.

Collection methods may include computer imaging, remote collections, or even assisted self-collection. Each method has its advantages and drawbacks regarding effort, cost, and completeness.

Weighing the pros and cons of every method falls on a lead investigator, who can seek advice from IT and legal. Balancing the needs of the investigation with the realities of budgets and scope can be challenging. Keeping an eye on proportionality and ensuring the ends justify the means, helps keep costs under control while ensuring the results are reliable and meaningful. Keeping the scope of the investigation in check will limit the amount of processing of sensitive personal data, something GDPR mandates.

Of course, there are trade-offs to each approach in terms of efforts, costs, and completeness. For example, computer imaging is more likely to be appropriate for cases involving suspected wrongdoing. This is the area where digital forensics enters the fray, since it provides opportunities to perform detailed searches of systems and recover potential evidence that has been hidden or removed. This applies to investigations into fraud, theft and interpersonal behavior such as harassment.

When relevant data is less likely to be hidden or removed, such as regulatory requests, general requests for information under privacy law, etc. self-collections and remote collection of data may suffice. In such cases, the eDiscovery solution can do the work without the need for the digital forensics toolkit.

## **Who should be involved in the collection?**

Investigators can either perform the collection of the ESI themselves, or assemble a collection team. For a collection team, investigators should aim to select employees who are up to the task of collecting materials internally, or assist custodians during the self-collection process. Employees involved in the collection activities should be familiar with data handling procedures as well as the case management tools, where the process of handling will be tracked.

Employees who handle ESI should update the case management tool at each and every transfer to create a proper chain of custody for the evidence. If these updates are not properly performed, the results of the investigation may be called into question. Even if the case in question doesn't involve the courts, a defensible collection process is important to meet the needs of regulators, or to ensure the results of an internal investigation are reliable.



## Chapter Four

# Data Retention & Preservation Policies

A data retention policy is a must-have for any organization that deals with data. For companies that are subject to external oversight or are active in litigation-sensitive or regulation-heavy fields, such a policy is essential. Retention policies focus on managing (from creation to destruction) the records necessary for the organization to conduct business. It follows that retention policies and procedures are utilized to create, store, use and discard business records.

Although data preservation is related to general data retention policies, it has a few significant differences. As said, a preservation period can arise at any time during a record life cycle (unlike a retention process, which always starts at the creation of a record).

Data preservation exempts the subject data from the regular data retention process. Simply put, a record cannot be destroyed until the pending or upcoming investigation is completed. Once it is, the preservation policy is lifted and the relevant retention policies once again apply.

Failure to properly preserve information and records may result in sanctions ranging from default judgments in civil cases, to monetary fines or even imprisonment in relation to governmental investigations. Data preservation ensures that all relevant data is preserved for the duration of the investigation (and any follow-up that may be needed). Retention policies, especially with regards to the destruction of records, do not (and cannot be allowed to) apply to any data subject to data preservation.



## Chapter Five

# Data Collection

Data Collection is the acquisition of potentially relevant ESI, as defined by the identification phase. In the context of litigation, regulatory inquiries, and internal investigations, this information (and its related metadata) needs to be collected in a justified, proportionate, efficient, and targeted way. As such, the collection process needs to be well-documented and defensible throughout. At the same time, as data is collected, its contents may provide feedback to the identification process, potentially impacting the scope of the overall process. This may also have an impact on the amount of data that needs to be collected.

### Effective Collection

The data collection process is usually both time-consuming and disruptive. However, there is little to no room for cutting corners: even after

collection, it is important to check the quality of the collected data as it is delivered. These checks are part of the validation stage of the collection process.

A common approach to validate the integrity of the data is to apply hashing to the original and copied data. These results can be compared afterwards: if the identifiers of the original and copied data match, the pieces of data are considered identical. If the validation process shows data is missing, incomplete, or incorrect, additional collection may be necessary.

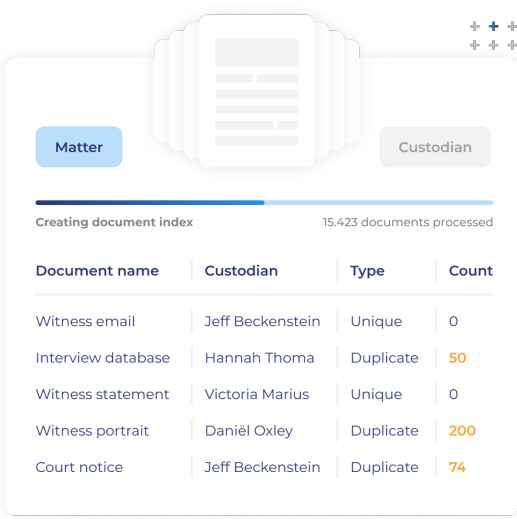
## Chapter Six

# Processing Collected ESI

Following the process of mapping, extracting and collecting the relevant ESI, eDiscovery solutions can be used to search the results. During the processing phase, data is analyzed and prepared for review. This is where data is culled (or cut down) prior to review. Unlike mapping and collection, processing isn't always an essential part of the investigation: Sometimes, especially during law enforcement investigations, the requestor asks for the raw data sets to be delivered. Even when this is the case, the raw data should still be processed by investigators, during the execution of the so-called shadow investigation which is held based on the knowledge provided to the investigation by the requestor.

Unless the requestor explicitly requests (or requires) data be unprocessed, processing it prior to handing off to reviewers helps save time, effort, and resources by reducing the amount of information to be reviewed. Tools need to be in place for data to be culled properly: specialized investigative tools are built for this purpose, and cull data in such a way that the processing itself is documented. This allows the parties that receive the results of the investigation, internal or external, to retrace the steps taken to cull data, so they can verify that the results are based on a sound dataset.

Processing isn't only culling, though. Information extraction is also an important function. This means creating (machine) readable data out of compound or non-searchable objects contained in the dataset. Compound objects include compressed files, imagine a ZIP file attached to an email in someone's inbox: that ZIP file can contain multiple separate files that may contain important information. An eDiscovery tool can unpack such files and check if there's anything important in there.



The screenshot shows a software interface for document processing. At the top, there are two buttons labeled 'Matter' and 'Custodian'. Below them is a progress bar labeled 'Creating document index' with the text '15,423 documents processed'. A table below the progress bar lists document details. The table has four columns: 'Document name', 'Custodian', 'Type', and 'Count'. The rows are: 'Witness email' (Jeff Beckenstein, Unique, 0), 'Interview database' (Hannah Thoma, Duplicate, 50), 'Witness statement' (Victoria Marius, Unique, 0), 'Witness portrait' (Daniël Oxley, Duplicate, 200), and 'Court notice' (Jeff Beckenstein, Duplicate, 74). The 'Count' column values are highlighted in orange. The interface is surrounded by a decorative border of small blue plus signs.

Document name	Custodian	Type	Count
Witness email	Jeff Beckenstein	Unique	0
Interview database	Hannah Thoma	Duplicate	50
Witness statement	Victoria Marius	Unique	0
Witness portrait	Daniël Oxley	Duplicate	200
Court notice	Jeff Beckenstein	Duplicate	74

Non-searchable objects may sound like something that doesn't come up frequently, but think about scanned receipts for example: those aren't readily searchable. In addition to scanned documents, PDF files, bitmaps, video/audio files, etc. are all non-searchable by default, and need processing to be made readable by a search tool. eDiscovery tools may use OCR (optical character recognition) and/or audio search to transcribe sound files. Through these methods, the software transforms the non-searchable into a format both human and machine reviewers can work with.


Once extracted, the culling stage of the processing can begin. At the very least, tools should ensure the data is without double entries through what's called deduplication. More advanced tools (such as [ZyLAB ONE](#)) enable users to cull data by using filters and queries, reducing the amount of data eligible for review.



# Reviewing and analyzing the information

Finally, after all that work to map, collect, and process the data involved in the investigation, it's time to start actually investigating. Reviewing and analyzing data are a package deal here: reviewing data simply means evaluating the data for relevance, while analyzing evaluates that relevant data for content and context.

into the custody of the investigation. Now that the dataset has arrived, reviewers need to prepare for their task, which is to establish relevance of the data in their set. Regardless of the size of the review team, two things need to be established prior to starting the review: the review strategy and the review environment



Although they perform the lion's share of the actual investigation, reviewers join the process relatively late.

For the purposes of this document, the focus will be on the review phase. Although analysis is an important part of the process as a whole, we'll not discuss it in-depth here: it is exceedingly difficult to standardize weighing context and content in a dataset. That process is defined by the content and context of the ESI, meaning there's no one-size-fits-all approach.

### Preparation and review strategy

Up to this point, most of the process as a whole has been about finding data and moving data

Establishing the review strategy means setting up the protocols that define how the review will be conducted, set up a timetable, and establish terminology to use for tags, codes and annotations. If the dataset contains some amount of foreign-language materials, the strategy should define if this should be left as-is, machine translated, or translated by humans. If tools permit, the usage of Technology Assisted Review (TAR) should be noted here as well. Finally, a protocol for handling sensitive, confidential, or privileged data should be put in place.





Once the review strategy is finished, the review environment needs to be prepared. Reviewers need to receive user access rights to the tools they need to perform their duties, and be given instructions and training.

### **Reviewing the dataset**

With the dataset, strategies and tools in place, the review can begin. While the data is weighed for relevance, detailed logs should be kept, so it can be included later during the presentation of the data as a technical report.

The low-tech way of reviewing is, as the name implies, low-tech. It mostly consists of reviewers sifting through the information manually and weighing documents for relevance. Predictably, this process is time-consuming, labor intensive, and leaves significant room for error. .

At the end of the day, reviewers are human. That means the droning, repetitive nature of reviewing

will eventually get to them, which invariably leads to mistakes, inconsistency, or concentration lapses. Also, since humans are only equipped with a single pair of hands and eyes, they're fairly limited in terms of how much data can go through their hands for them to see. This means low-tech manual review tends to take a lot of time.

The high-tech way of reviewing means using advanced investigative tools to help mitigate the limitations of human reviewers. Modern eDiscovery solutions have advanced tools to quickly and automatically cull irrelevant data from a set using Technology Assisted Review (TAR), a process through which a reviewer 'trains' the solution, powered by Artificial Intelligence (AI) to tell the difference between relevant and irrelevant data. Once the AI understands the difference, it can classify documents based on input from reviewers, in order to expedite the organization and prioritization of the dataset.

Technology Assisted Review can dramatically cut down the time (and cost) of reviewing, as reviewers now only need to review a dataset pre-selected for relevance. Of course, the input process for the AI's training set will be documented in order to preserve defensibility. It's important to note that TAR doesn't mean human reviewers are not involved at all; verification remains important – the A in TAR stands for Assisted, not Autonomous.

Regardless of which method of review is used, the end of the review phase yields a culled dataset composed of only relevant material.

### Analyzing the dataset

Whether or not review is the final step depends on the context of the investigation. If it concerns an external request, the review dataset moves directly to the presentation phase. For internal investigations, the review dataset will need to be analyzed as well. If the information request originates externally, the analysis will be performed by the requestor.

No matter who does the analysis, this part of the process may create a feedback loop: if analysis shows that the dataset is missing information, the review process starts up again to provide it.

If the missing information is not present in the review dataset, this information can be sought outside of it. The combined model of digital forensics and eDiscovery allows for this through the analysis feedback loop (see Fig. 5). This is done by making use of the data recovery abilities of digital forensics tools if the data is lost, or by entering the information into the eDiscovery solution through the conventional means of data collection.

The goal is to find key patterns, topics, people and conversations. While the review answers the question: 'Where is the relevant information?', analysis answers: 'What is the relevant information?'.

We won't delve too deep into analysis here, but suffice to say that modern end-to-end eDiscovery solutions offer a wide range of visualization options and analytical tools to identify and show the connections and content within a dataset.

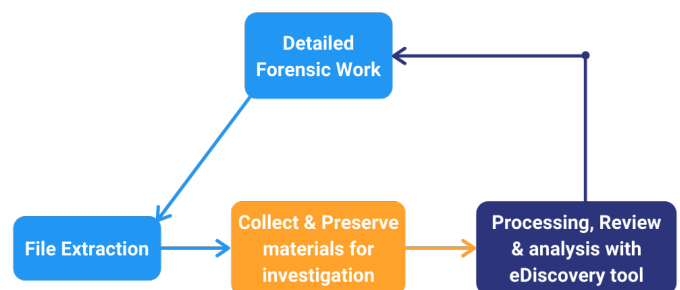


Fig. 5 - Analysis feedback loop (detail)



## Chapter Eight

# Producing the results

Once the review dataset is believed by investigators to be complete, the results of the investigation will need to be prepared for either internal or external analysis. The most important decision at this stage is determining the production format.

A few options are available:

**Native format:** files are produced as they were originally – authentic, but difficult to redact;

**Image formats:** files are reproduced into an image format, such as PDF or TIFF; - this is the most common output format used by eDiscovery solutions. TIFF allows for the redaction of

information, the embossment of Bates numbers as well as other information right on the document. Image files can also be opened on any system without the need for access to the original application used to create it.

**Metadata or Load Files:** files of this type are typically produced to provide the metadata of documents in the production dataset and often the tagging work product performed within the solution. These files can take a variety of industry-standard formats.

**Extracted Text:** this is the full text of the documents separated from its original format. This can be printed, too.

## Delivering the investigation

Once produced, the dataset should be delivered.

If that delivery is external, or concerns highly sensitive information, the security of the dataset while in transit must be the primary concern.

These concerns can be taken away in a variety of ways: by making use of secure file sharing services or physical data carriers to perform the transfer.

Furthermore, encrypting the files prior to transfer can bolster security. Perhaps the best way is to sidestep the issue, allowing the presentation of the results to be done in the eDiscovery platform - that way there's no transit at all .

In addition to the results of the investigation, a technical report of the processes may be asked for. Depending on the type of investigation, the technical report may include the following:

- The data map used for identification of the ESI. This data map can also provide

event tracking for the process from data identification to the upload to the eDiscovery solution.

- A report of data that could not be collected (and why);
- An event log of actions performed during processing;
- A log of the operations applied to cull the dataset;
- A copy of the review strategy document;
- A log of the data produced and formats used;
- Hashes of the produced data.

Once the production data and technical report have been handed over, the digital investigation is completed. The data within the eDiscovery solution can be archived according to the case retention period set in the retention policies.





## Conclusion

# End-to-End Digital Investigations

By using a combination of tools for both digital forensic work and eDiscovery, organizations can be sure that their ability to self-investigate remains intact. As the data environment modern businesses operate in become increasingly diverse and complex, as does the need for robust tools capable of processing the large amounts and variety of ESI in these environments.

Having such tools, and people able to use them, available is essential for any organization that finds itself needing to comply with internal and/or external investigatory requirements.

As noted, most of the tools and applications of eDiscovery tools are based on civil litigation procedures in the United States. When it comes to performing comprehensive digital investigations, the reference models typically associated with eDiscovery have to be adapted to this context.

Specifically, the addition of avenues for gathering information that isn't readily available for collection - in particular the addition of digital forensics as a potential source of information. This addition, the necessity of which is mostly due to in the different nature of the use case for eDiscovery outside the United States, shows where eDiscovery fits in terms of an end-to-end digital investigation.

In spite of the models needing a few tweaks, it is undeniable that its growing importance in American civil procedures has been a great boon to the development of eDiscovery technology. Outside of the courtroom, these developments are still incredibly useful for those who need to investigate the growing data stores.

Most notably, the integration of AI and text mining techniques into eDiscovery tools make them uniquely suited to improve the speed, thoroughness and accuracy of investigations. By eliminating most of the highly repetitive and boring tasks surrounding the initial review, eDiscovery tools help keep investigators' eyes fresh when it comes to reviewing potential evidence. Meanwhile, these initial sorting of documents, the elimination of irrelevant information and/or duplicates happens significantly faster once the AI is underway.

To see an eDiscovery solution in action, or ask any question you have regarding the EDRM and its applicability for internal investigations, we're happy to show you around our solution. Request a demo [here](#).



an IPRO company

The eDiscovery platform for legal fact-finding. Trusted by government, law firms and corporations.

**Why ZyLAB?** →

#### **Location**

Laarderhoogtweg 25  
1101 EB Amsterdam  
The Netherlands

#### **Contact**

[www.zylab.com](http://www.zylab.com)  
[info@zylab.com](mailto:info@zylab.com)  
+31-20-7176500

---

#### **© ZyLAB Technologies B.V. and/or its affiliates ("ZyLAB").**

No part of any ZyLAB blog, whitepaper, datasheet or any other marketing publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of ZyLAB. The information, data and content contained in such ZyLAB marketing publications is owned by ZyLAB and is subject to change without notice. ZyLAB assumes no responsibility for any errors that may appear.

All ZyLAB's marketing publications are educational in nature and are not legal advice for an organization's particular circumstance. Marketing publications are for informational purposes only. ZyLAB makes no warranties, expressed or implied, by operation of law or otherwise, relating to these documents, the products or the computer software programs described herein.

#### **ZYLAB DISCLAIMS ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.**

In no event shall ZyLAB be liable for (a) incidental, indirect, special, or consequential damages or (b) any damages whatsoever resulting from the loss of use, data or profits, arising out of these documents, even if advised of the possibility of such damages.