The Legal Intelligencer

Investigations Are Not Discovery, So Don’t Treat Them as Such

By Tara Lawler and Leonard Leonard Impagliazzo

Tara Lawler, left, and Leonard Impagliazzo,
right, of Morgan Lewis & Bockius. Courtesy photos

June 22, 2022

In this installment of our continuing series on e-discovery basics, we compare discovery strategy for internal investigations with traditional e-discovery. Given the need to quickly determine whether laws, regulations, or internal company policies may have been violated in an internal investigation, we will discuss the importance of successfully leveraging technology to identify any potential wrongdoing to help mitigate criminal prosecution, civil fines, restitution, or reputational risk in a quick and efficient manner. In our last article, we discussed the importance of developing a thorough understanding of client data and drafting a comprehensive discovery playbook in advance of litigation. Although concepts and practices relevant in e-discovery are useful in conducting an internal investigation and there are similarities between both processes, discovery for litigation and “discovery” for internal investigations are markedly different and should be treated that way. This article will explore the process by which a skilled eData attorney can carry out an internal investigation to completion.

Internal allegations are conducted when an organization becomes aware of allegations of potential wrongdoing either from outside the organization (e.g., regulatory request, media) or from inside the organization (e.g., whistleblower, company “hotline,” audit). Whether the allegations involve insider trading, fraud, misappropriation of funds, sexual harassment, or any other type of malfeasance, the risk to the organization can be immense and should be thoroughly investigated.

The eData attorney is uniquely positioned to guide an organization through a diligent and efficient investigation of potential wrongdoings. The eData attorney possesses a unique skill set, including expertise in information governance and preservation best practices, mastery of technology, and deep experience in a vast array of subject matters.

Traditionally, internal investigations have been handled by a team of lawyers that collects a relevant set of data and runs broad and simple keyword searches. Based on these searches the team then conducts a linear review of thousands of documents, taking notes and manually compiling summaries throughout the process. Although keyword searches and linear reviews are still common in litigation, this is not an effective or efficient process for conducting an internal investigation.

Internal investigations lack the defined scope of traditional discovery and are not bound by federal or state rules of civil procedure. In traditional discovery, the eData attorney has the benefit of identifying data based on a factual understanding of the allegations and specific requests for production. Internal investigations are undefined and potentially expansive. Investigations and traditional discovery do, however, share some of the same phases found in the electronic discovery reference model (EDRM): identification of potential sources of data, preservation, and collection, as well as processing and review. However, the internal investigation is unburdened by production requirements, privilege logging, and other external factors. The investigation is a prime opportunity for the eData attorney to leverage the full suite of available technology solutions while finding synergies among technology, processes and workflows.

The first step in an investigation is to identify the alleged wrongdoer(s) and other employees whose data may also need to be preserved and collected. For example, if the potential wrongdoer is a sales representative who is alleged to be conducting off-labeling marketing, the organization may want to preserve and collect the alleged wrongdoer’s data as well as data belonging to others within his or her territory or product line to determine if other employees are making similar violations. These decisions typically depend on the allegations and the source of the allegations (i.e., internal source versus government inquiry). The eData attorney has experience making these strategic decisions and can guide the organization on best practices.

After identifying the alleged wrongdoer and other employees, the eData attorney will work with the organization to set the proper preservation strategy. The target(s) of an investigation may have an incentive to dispose of evidence related to his or her misconduct, or the organization’s regular retention practices may lead to the disposition of key information. It is important to preserve information at the earliest opportunity.

Unlike the traditional e-discovery process, where a legal hold is issued to inform the recipients of the hold to preserve all potentially responsive information, in an internal investigation the targets of the investigation are typically not informed of the investigation to avoid potential spoliation. In the context of an internal investigation, we refer to these targets as “silent custodians” under a “silent” legal hold. The “silent” legal hold serves as a preservation and tracking mechanism to preserve the targets’ data on the back end.

The target(s) should not be made aware that his or her data will be collected during the investigation. The collections should be executed on the “back-end” if possible, to avoid potential manipulation or deletion of data by the alleged wrongdoer. This has some drawbacks in that it impairs the ability to collect potential key information that the organization would typically need assistance from an employee to collect, such as data residing on an alleged wrongdoer’s mobile device(s) or third-party communication applications. Nonetheless, the eData attorney will begin the investigation by collecting information from data sources that reside on the organization's servers (e.g., email, OneDrive, shared drive) and that do not require involvement from the alleged wrongdoer. If wrongdoing is uncovered in these company-based data sources, then a decision can be made to collect additional sources, such as mobile device data, that require informing the alleged wrongdoer of the investigation.

Another drawback of the back-end "silent" collection is that it prevents an organization from conducting targeted collections. In most traditional e-discovery matters, targeted custodial collection interviews are conducted to reduce the volume of data collected. Internal investigations, however, typically require full collections of the alleged wrongdoer's data unless a date restriction can be applied. The eData attorney can guide the organization on the decision of whether to use a shorter timeframe (thereby reducing the volume of data collected), which runs the risk of missing key information if the date range is not broad enough, or to collect a wider date range that will increase the volume of data collected but will eliminate the need to re-collect if the data range is incorrect. This decision depends on the severity of the allegations made and the urgency to identify any potential wrongdoing.

Data that is collected in a nontargeted manner (even if date ranges are applied) often results in intimidating volumes of data that must be analyzed in a short period during an internal investigation. It is imperative that the eData attorney work collaboratively with the organization and merits counsel managing the investigation to develop a working understanding of the known facts related to the allegations. Once the data is collected, the eData attorney can add tremendous value by leveraging technology to identify a relevant rich data set from within the broad collection. The eData attorney will employ discovery tools, such as Relativity, to leverage all analytic tools at his or her disposal. Analytic tools employ artificial intelligence and machine learning to allow an eData attorney to gather insight into a dataset while leveraging various inputs and outputs to identify consequential associations between documents. Using the analytic tools, the eData attorney can quickly identify datasets that are pertinent to the alleged wrongdoings by reducing the volume of data requiring analysis. This filtered data set is identified after applying global de-duplication, thread suppression, and further document-level deduplication.

The eData attorney will then conduct a communications analysis of the filtered data set to identify patterns of communications (e.g., dates, times, words). The communications analysis allows an eData attorney to visually observe clusters and webs of incoming and outgoing communications, as well as strategically focus on patterns, if any, of internal and external email domains. The communications analysis allows the eData attorney to leverage technology to quickly identify communications patterns of interest over an extensive dataset. Some of the goals of this analysis should be to identify and capture any email aliases utilized by the target(s), as well as call out irregular communications patterns and key terms. Depending on the allegations, the eData attorney may also focus this analysis on the presence of third-party domains and assess whether, for example, proprietary information is being circulated outside of the organization.

It is at this stage where the collaborative process comes deeply into play. The eData attorney works closely with the organization and merits counsel to review the results of the communications analysis to develop a culling strategy. The purpose of employing a culling strategy is to further narrow the data collected to a relevant rich dataset. This culling strategy includes the identification of the individuals, organizations, and key dates to be targeted within the relevant rich dataset. The eData attorney will leverage prior experience on similar investigations to develop sophisticated matter-specific search terms. These search terms will help pinpoint potential wrongdoing (or lack thereof) and remove from the dataset "junk" or irrelevant information.

The culling strategy can be bolstered if the organization has already identified key documents outside of the review process—for example, if documents were provided or referenced in the government inquiry or if a whistleblower provided materials. These key documents can be leveraged to generate a technology-assisted review (TAR model) to identify documents within the larger dataset that have similar content to the key documents. The eData attorney will execute the TAR model to identify the documents with the highest likelihood of containing similar content and review those documents for substance first. While the TAR model is utilized to identify relevant rich data, human review is necessary to differentiate between relevant and nonrelevant content that may share textual or conceptual similarities. Throughout the review, the identification of both relevant and nonrelevant material is crucial in further refining the TAR model and the associated relevant rich dataset.

While the goal of the internal investigation is to identify the existence or absence of wrongdoing, the identification of nonrelevant data does present value. Once nonrelevant data is identified and validated, coding can be pushed out to remove duplicates and lesser-included threads from the review universe. This would allow these similar documents to be removed from the review population as well as further bolster the TAR model. The TAR model performs better with samples of both relevant and nonrelevant data.

After identifying additional key documents, the eData attorney will target family members (e.g., attachments to emails) of the key documents and then update the TAR model to include the key family members. This is an iterative process and will be repeated until all relevant content is identified. The eData attorney will then assess the relevant data to further bolster the initial communications analysis and make further refinements to the search terms if appropriate. Once the investigation concludes, and if the organization requests it, the eData attorney can provide a report to the organization and merits counsel summarizing the data collected and the analysis taken and identifying key documents. This represents an additional opportunity to utilize additional Relativity applications, such as Case Metrics, that help generate chronologies and timelines. Throughout this process the eData attorney should document the steps taken to arrive at the relevant rich dataset.

In our next installment of this series, we will focus on some of the pitfalls and pain points of privilege logging and developing best practices for dealing with privilege logging.

Morgan Lewis & Bockius partner **Tara Lawler***'s practice focuses on e-discovery, information management and data privacy. Associate* **Leonard Impagliazzo***'s practice also focuses on e-discovery, information management and data privacy.*