

# Digital Forensics & eDiscovery

TBAA Report

Discussion of a Potential Forensic Examination  
Scenario Highlighting Key Aspects in the Subject

M.Sc. Cybersecurity  
School of Computing  
National College of Ireland  
Dublin, Ireland



Saptarshi Laha  
x18170081

# Contents

|  |    |
|--|----|
| Basic Information.....   | 1  |
| Executive Summary.....   | 1  |
| Case Background.....   | 1  |
| Case Acceptance.....   | 2  |
| About the Investigator.....  | 2  |
| About the Digital Forensics Examiner.....  | 2  |
| Scope of the Investigation.....  | 3  |
| Case Summary and Objectives.....   | 3  |
| Summary of Conclusions.....  | 3  |
| Answer 1 (A).....  | 3  |
| The Digital Forensics Investigation Methodology.....   | 3  |
| Answer 1 (B).....  | 7  |
| Effect on Digital Evidence Due to Potential Tampering.....   | 7  |
| Answer 1 (C).....  | 8  |
| The Methodology of Analysis and Discovery from Digital Artefacts and the<br>Uncovering Potential Digital Evidence from them..... | 8  |
| Answer 2 (A).....  | 13 |
| The Electronic Discovery Reference Model.....  | 13 |
| Technology-Assisted Review.....  | 18 |
| References.....  | 23 |

## Basic Information

**Table 1.** General Information regarding the investigation

|                                   |   |
|-----------------------------------|---|
| <b>Investigator</b>               | Mr. Mark Monaghan (Professor)   |
| <b>Digital Forensics Examiner</b> | Mr. Saptarshi Laha (Student)  |
| <b>Accused</b>                    | <b>Peter Banks</b> and <b>Mary White</b> (Former employees of a high-quality furniture production company, Top Office Furniture Limited, and, currently directors of the recently established printing company, Full Bleed Limited) |
| <b>Offence</b>                    | Exfiltration of Client related sensitive data   |
| <b>Date of Request</b>            | 26 <sup>th</sup> April, 2020  |
| <b>Date of Conclusion</b>         | 19 <sup>th</sup> May, 2020  |
| <b>Report Publish Date</b>        | 19 <sup>th</sup> May, 2020  |

**Disclaimer:** The chosen case scenario is for learning purposes only, and any association to an actual case and litigation is purely coincidental. Evidence presented in the case scenario is fictitious, and the intention is not to reflect any actual evidence. Reference herein to any specific commercial products, processes, or services by trade name, trademark, manufacturer, or otherwise does not constitute or imply its endorsement, recommendation, or favouring by the Council of the European Union, and the information and statements shall not find its use for advertisement of any sorts. The information and views set out in this report are those of the author and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use of any information contained therein.

## Executive Summary

### Case Background

The CEO (Chief Executive Officer) of Top Office Furniture Limited fears that Peter Banks and Mary White are involved in the exfiltration of sensitive information based on their recent resignation notices that got handed to the CEO. Peter Banks, the most experienced employee of the organisation, was the first to hand in the notice of resignation on the grounds of starting a restaurant with his wife, and, Mary White, the second most experienced employee of the organisation, handed in the notice of resignation within one day's gap of Peter Banks, resigning on the grounds of a career change to work with a start-up company. On further investigation by the CEO, he found that a new printing company was registered and the directors of the same were his former employees that resigned. Having known this, the CEO decided to put his employees serving the notice period on gardening leave, after reclaiming the devices provided by the office from them. After the duration of Peter's term of the

contract (as mentioned by the scenario “After Peter left the organisation.” Suggesting that the term ended) the CEO was able to discover traces of evidence that led him to believe that there might have been a case of exfiltration of data to win business from his company Top Office Furniture Limited in upcoming contract negotiations and potentially undercut him on current contracts.

## Case Acceptance

Forensics Pvt. Ltd. has undertaken the task of carrying out the digital investigation of the digital artefacts related to the case. The date of the request to carry out the digital investigation was the 26<sup>th</sup> of April, 2020. Forensics Pvt. Ltd. was given time until the 19<sup>th</sup> of May, 2020 to report the findings based on the investigation. Mr. Mark Monaghan, the lead investigator of the case, has accepted the case and further assigned the digital forensics examiner, Mr. Saptarshi Laha, for carrying out a detailed investigation of the digital artefacts under his guidance while maintaining an appropriate chain of custody.

## About the Investigator

The head of Forensics Pvt. Ltd., Mr. Mark Monaghan, is a highly proficient individual in the field of digital forensics and e-discovery and his primary task is to act as a link between his subordinates and the client. The task of the Investigator, in this case, additionally includes guiding the digital forensics examiner in structuring the uncovered data and conveying the updates to the client at regular intervals regarding the progress. The Investigator also acts as a point of connection between the legal team and the technical team of Forensics Pvt. Ltd. to help build a strong case based on the findings of an incident.

## About the Digital Forensics Examiner

The digital forensics examiner assigned to this case, Mr. Saptarshi Laha, is responsible for performing the forensic analysis of the following digital artefacts present on the crime scene

—

- 1. Suspects’ Landlines.**
- 2. Suspects’ Desktop Computers.**
- 3. Suspects’ Work Email Accounts.**
- 4. The Central Server.**
- 5. The Email Server that has Exchange installed.**
- 6. The Printers present in the office.**
- 7. The Internal Network of the office.**
- 8. The Printer Server storing Print Logs.**
- 9. USB sticks and DVDs found.**
- 10. The Mobile Phone found.**

Additionally, the digital forensics examiner’s role, in this case, entails the reporting of facts uncovered from the investigation of these devices and handling this evidence with care and caution to aid the formation of a strong case with the help of the legal team, if necessary, under the guidance of the Investigator.

## Scope of the Investigation

The scope of this investigation entails the finding of relevant pieces of evidence regarding any exfiltration of data or other illegal or demoralising activity performed by the two former employees of the organisation. Forensics Pvt. Ltd. got called into the scene after the CEO performed some of his investigations manually, which may have altered potential evidence. Despite maintaining a proper chain of custody, the interference caused by the CEO of the company in the process might produce results that may drastically differ from the ideal and unaltered results.

## Case Summary and Objectives

The objectives for the case defined by Forensics Pvt. Ltd. is to:

- 1. Finding every bit of evidence from all the digital artefacts that are part of the crime scene and trying to link it with the former employees of the company where applicable.**
- 2. Create a report detailing the process of uncovering relevant facts to aid the reproduction of the same in the court if necessary by an external digital forensics examiner.**

The summary of this case entails the process of performing a digital forensics examination from a contract standpoint for an organisation, assuming there has been no tampering or mishandling of data or artefacts, which is clearly not the case and the consequences are discussed later as a part of an answer.

## Summary of Conclusions

The conclusions could only get highlighted if the actual evidence material got presented as a part of the scenario. However, since only the scenario got presented, hence only guidelines could be highlighted that needs following for building a strong and successful case, in the case of a breach in sensitive information as the CEO doubts.

## Answer 1 (A)

### The Digital Forensics Investigation Methodology

The steps for carrying out digital forensics investigation in a professional setting are as mentioned below. Additionally, the steps present below also highlight the different activities that should get performed by the digital forensics examiner in this case. In general, digital forensics is more of an art than science. As every case is different, the methodology useful for performing active digital investigation is different in different situations. However, a generalised overview gets presented, taking into consideration the scenario to which the digital forensics examiner gets summoned.

- 1. Preservation** – Preservation of the crime scene is one of the essential tasks that a digital forensics examiner tasked with carrying out an investigation has to perform. In the case of this incident, the preservation of digital evidence might be very challenging, depending on the notice period that Peter served. This problem arises as the CEO of the company was waiting until the contract term of Peter ended before performing his analysis and further deciding to call in the forensics experts.

Additionally, the fact that the evidence has been tampered with by the CEO by multiple different activities mentioned in the next section (**Answer 1 (B)**) makes it even more challenging to gather digital evidence while maintaining a proper chain of custody.

The unnecessary downtime before calling in the digital forensics team (while waiting for Peter to serve his notice period) is a critical issue because information present in the memory may have easily gotten overwritten or wiped. Apart from that, the hard disk sectors that contained deleted file pointers may have gotten overwritten as well due to the addition of new files by external factors or as housekeeping tasks were performed by the operating system installed. Multiple registry keys might have gotten modified due to this, such as potential timestamps of applications that got executed, which would help in the building of a case timeline or settings related to previously present applications that got deleted in the case of Mary's computer system. Taking all of this into consideration, along with the possibility of the fact that the worst-case scenario gets presented to the forensics investigator, it is rather tricky to acquire potentially useful data from the systems in general.

However, some systems are far more accessible and preserve their data for a more extended period in terms of evidence than others because external logging takes place, and the employees do not get endowed with complete control of the same. This fact can easily get demonstrated by the example that the company can acquire logs of telephone calls and mobile phone calls that got placed, and the same information found thereof can get leveraged to find other useful information on the system related to the same and the people called to further connect the dots in this case. This method of evidence collection is only possible as the network provider or an external agency is responsible for keeping the records. The same can get performed for inhouse servers such as the printer server assuming it did not get left in the open unlike the email server or the central server and the employees were not competent enough to utilise the functionality of the central server to alter the records of the other networking devices present. It also depends on the fact that if the employees had escalated privileges on the server or not, because, both of the suspects were senior employees of the organisation and could easily have such sensitive information trusted to them by the CEO or the higher management.

Additional complications arise from the fact that USB devices and DVDs could easily get plugged in, or the unmonitored internet could easily be misused, and applications or other such digital artefacts could get executed, making matters even worse. One would, in that scenario, not even have to be tech-savvy, and instead, just follow a set of videos or guides available online to perform something malicious which could easily be commonplace in this scenario as is the case for Mary's system where a potential disk wiping solution gets discovered by the CEO, while it underwent execution. Also, depending on the complexity of this program, standard methods of data recovery would not be possible from the disk, and one would have to rely on sophisticated

hardware and methods such as magnetic disk microscopy to extract potential evidential data.

2. **Acquisition** – In this step, data gets collected from the artefacts discovered that undergo preservation in the previous section. Since the preservation step was not in the ideal or perfect state that is desired by any forensics examiner, the results of the acquisition might drastically differ, or in the worst-case scenario, it might not even be possible to gather. In this phase, photos should get taken of the running programs on the systems present, apart from taking photos of every other acquisition made in the crime scene to maintain a proper chain of custody. Additionally, a volatile and non-volatile memory snapshot (which is nothing but a bit for bit copy of the device) should get taken for the suspect systems using DMA controllers for volatile memory and write-blockers for non-volatile memory for exact replication of the crime scene in the lab environment. Every piece of original evidence should then get placed in separate evidence collection bags with the tags mentioning the date and time of acquisition of the data, apart from its location and the state in which it got discovered. As a general rule of thumb, the more information that is present about items present in the crime scene during the acquisition process, the better the case is and the more reliable the documentation will be to aid the case.

Acquisition of the other information related to the central server, the exchange email server and the server logs can similarly take place, and the volatile and non-volatile memory components apart from the system logs should get exported for analysis in lab to a different device as it is incredibly unideal to perform any analysis on the actual device. Additionally, the mobile phone should be taken into custody, apart from the records of incoming and outgoing calls of the landline and a memory snapshot of the same if the landline supported smart features to help build a strong case with their respective tags and timestamps. The mobile should also be turned to flight mode to prevent any erasure of data by the suspects or tampering of evidence caused by unintended activities such as incoming calls or text messages. The internet should also get disconnected from all of the devices that are supposed to undergo digital forensics examination as the suspect could remote wipe the devices, thereby leaving no trace and making extraction of potential evidential data related to the crime even harder. Also, turning off the devices if they are in a turned-on state is generally considered a bad idea as the memory information could get lost and a memory dump could not be possible. The comprehensive set of tasks performed by operating systems during a shutdown also known as housekeeping tasks might lead to the erasure of potential evidence related information from the hard disk drives and thus turning off the system is generally discouraged.

In this case, since the central server got left in the open, which also contained the email server, apart from the ease of access to suspects' systems via USB sticks, DVDs or unmonitored internet connection makes it incredibly challenging for the forensics examiner to acquire useful data. The delay in performing the acquisition and the

waiting time also plays a vital role in the potential loss of digital evidence. The same is the case because of the tampering of the evidence performed by the CEO of the company.

The only conclusive evidence that can get acquired in this case is from devices that underwent external logging such as mobile phone records, landline call records, and the printer records assuming that the print server was not compromised.

3. **Analysis** – The steps concerning the analysis and the discovery of various pieces of evidence get discussed in detail under the **Answer 1 (C)** section of this report. However, this phase entails a thorough analysis of the evidence collected in the previous phase to build a timeline of activities and understand the activities performed by the suspects. This phase includes the analysis of every single digital component in separate ways specific to the type of device encountered for potential information relevant or irrelevant to the case. The analysis always gets performed on replications of the original acquisition rather than the original acquisitions themselves. These replications should be such that, one could revert to the original state after performing specific alterations within a current image in a particular workspace, and hence snapshots are the most suitable candidates for the same. This methodology is essential to the workflow as it guarantees the preservation of the original digital artefact in its actual state so that it could get produced as evidence to the court, if necessary. However, if the same does not get performed, then there is a risk of potential tampering of evidence and dismissal as valid evidence by the court, thereby weakening the case.
4. **Discovery** – Discovery phase deals with the segregation of all of the data discovered in the previous phase and its categorisation into relevant or irrelevant data based on the constraints of the case. The operating system installed, usage of the system involved, automated tasks getting executed, etc. play a massive role in determining the relevant and irrelevant aspects in building a timeline of the suspects' suspicious activity linking them to the case. Once again, this phase is only possible and successful if the analysis phase results in acquiring of sufficient amount of data that can undergo categorisation or segregation based on the type of activity performed and the hypothesising on the reason why it was performed based on other conclusive pieces of evidence present to solve the puzzle. Discovery phase related to this case has also undergone an explanation in detail in the **Answer 1 (C)** section of this report, and hence just an overview of the phase is presented here.
5. **Documentation** – Documentation phase revolves around the production of a report based on the findings present in the analysis and discovery sections mentioned above. This documentation should be extremely extensive and should highlight every possible detail related to the case. One additional parameter the digital forensics examiner needs to keep in mind when generating or crafting this document by hand is to explain every step in enough detail so the results can get replicated and the same



output can be acquired in every case if it gets performed by another person. Another vital point to keep in mind is the technical jargon used in the report or the document generated should get mellowed down so that the general audience can understand and interpret it. This step is essential as the judge undertaking the ruling for the case in most of the circumstances will not be technically sound to understand the in-depth technicalities mentioned if the same gets presented in the documentation provided for the case.

Additionally, an extensive amount of care should get taken in building this document as this document serves as the deciding factor regarding the winning or losing of the case. Irrespective of ground-breaking discovery and analysis phases that get performed by the forensics experts, the document produced, if not detailed enough, cannot articulate the findings to the general audience including the judge and thus might lead to the organisation or the individual losing the case. Thus this is a further emphasis on the reason why the documentation phase probably plays the most critical role of the entirety of a digital forensics examination.

6. **Presentation** – Presentation phase deals with the usage of the documentation generated or created in the previous phase for building a strong case and taking it to court. Presentation phase deals mostly with how the data presented in the documentation provided gets interpreted, if it can get replicated and what it implies so that a flow of conclusive evidence based on the report can be presented in court to sue the suspect if they are found guilty. In this case, it is the task of the Investigator, Mr. Mark Monaghan, and the CEO of the company, to decide based on the evidence presented in the document if any further legal action is necessary. In case the same is essential, the Investigator will present the documentation to the legal counsel of Forensics Pvt. Ltd. and discuss with them the possibilities of a case and the grounds on which the suspect can potentially get sued.

The structure of the report has undergone division into two parts. The first part details the potential problems that could arise if there is potential tampering of evidence as present in this scenario and then discussing efficient and effective methods of relevant data analysis and discovery from each artefact encountered, and the best practices that should get followed in each of these circumstances. The second part of the report highlights the answers to one crucial question that is asked by the CEO in regard to the e-discovery procedures followed **(Answer 2)**.

### Answer 1 (B)

#### Effect on Digital Evidence Due to Potential Tampering

In this scenario, one of the most destructive actions taken against any forensics examination has gotten performed by the CEO. As digital forensics entirely relies on producing a chain of custody which deals with the handling of data and evidence, every system and digital components that are part of the crime scene and are subject to examination are be left in its initial state. Failing to do so leads to the failure in compliance with the requirements of a

digital investigation scenario apart from leading to potential intended or unintended tampering of data. This example could easily get explained by the fact that the USB present in the crime scene got inserted by the CEO to his personal computer for analysing its contents which when analysed in detail by the digital forensics examiner could lead to the linking of the CEO to the case, which is extraordinarily unideal and deviates from the truth. However, such an anomaly cannot generally get explained when providing evidence in the court and raises unintended questions in the mind of the judge, thereby making the case weak. Also, in the best-case scenario, this action would not alter anything, but in the worst-case scenario, it could delete the potential evidence that is present in terms of the Excel Spreadsheets or even introduces malware into the computer, further complicating the process of digital forensics.

Next, it is unclear how the CEO was able to gain access to the mails of one of the potential suspects. However, if the same gets directly accessed through the Exchange mail server (assuming that is the case as the enclosure to the same is kept open) then the applications recently executed on the server gets altered apart from the recently accessed emails, the recently performed activities and much more to name a few. Also, repeatedly performing such tasks can delete evidence in terms of deleted files whose file pointers present in the MFT (Master File Table) can get overwritten, rendering the discovery of deleted files even harder. The deletion of the file pointers is not problematic only in the case when active logging is present without any upper limit defining the number of records that get stored. This condition is scarce in this case due to the small size of the company and presence of an in-house server instead of a cloud-based one, and even then, the names of the deleted files could probably get logged, but it will still be highly unlikely that the actual contents of the file will get preserved, which still ends up making matters extremely complicated for the digital forensics examiner.

Lastly, the same holds if the CEO logs onto the desktops of the suspects as he did in this case, which could lead to the erasure of contents remotely by the suspects in the worst-case scenario, or set off an automatic disk cleaning program or overwrite file pointers to deleted files, or even modify other crucial aspects related to the operating system. Thus, the general suggestion always involves leaving the systems as they are and calling the forensics team at the earliest so that a proper chain of custody can get maintained in the entire process which can lead to the building of a strong case apart from making it possible for the digital investigators to gather as much conclusive evidence as possible.

### Answer 1 (C)

#### The Methodology of Analysis and Discovery from Digital Artefacts and the Uncovering Potential Digital Evidence from them

Potential data that gets observed and can serve as pieces of evidence in this particular scenario differ from system to system, and the methodology for performing an extensive analysis of the same gets explained depending on the system or digital component encountered in the list below –

1. **Landlines** – These can provide data regarding outgoing or incoming calls that got made from both the call logs present on the landline or from the network provider as

the company pays the bills and the registration of the phone is under the name of the company, and hence it is easy to request the network service provider for the same. Additionally, if the landline has a smart feature, the same can get dumped, or a snapshot of its operating environment can get taken for the analysis and discovery of other potential pieces of evidence such as text messages, saved contact information, and other potential information based on the type of smart functionality provided by the handset. The data collected from these processes can then get leveraged and cross-checked against the records of the clients for the discovery of conclusive pieces of evidence based on any anomaly discovered.

2. **Desktops** – Desktops serve as crucial evidence hotspots as they not only give a vast amount of information regarding the suspects' day to day activity but also makes the suspicious activity stand out in the timeline that gets generated due to the random occurrence of the same compared to the usage of other programs. This fact can easily get explained with an example related to the scenario where one of the suspects generally works on Microsoft Word but has recently been using a disk wiping utility which would make this activity stand out from their day to day routine.

Right places to look for in desktops for potential evidence include but is not limited to the recent files, the Windows Timeline in case of a Windows system, the registry hives (especially SAM, SOFTWARE and SYSTEM hives), the event logs of the operating system installed, searching the MFT for files that can get carved, looking for orphan files in the filesystem, analysing memory artefacts that get dumped, analysing all the files present on the operating system, etc.

Each of the places mentioned above leads to the potential discovery of a unique variety of evidential data. The recent files, for example, provides a brief overview of the recent activities that were performed by the suspects, in terms of accessing documents or other such content. The Windows Timeline feature on the other hand if enabled provides an overview of the recent activities that have been performed by the suspects including the execution of applications, accessing of documents and performing of data manipulation tasks with their respective timestamps.

Registry hives provide a whole lot of information regarding the operating system in general, such as the install date and time of the operating system, the registered owner of the operating systems, the list of installed applications on the operating system, the list of installed updated on the operating system, the list of potential users present on the operating systems, the media devices that were plugged in or used on the operating system, results related to the ShimCache and Amcache which provide a list of last few executed applications which help in the building of a timeline of the activities that got performed by the suspects; the web history, downloads and web plugins installed and other such crucial information.

The various event logs present on a system provide a plethora of information starting from the system turn on and off times, the number of successful logins to a particular account, even with the number of failed attempts to log in to a particular account. It also includes information regarding the executed application, warnings and errors recorded by the computer system including abrupt shutdowns and much more information which helps in building a strong case. Additionally, the event logs are much more detailed than the registry in maintaining a list of last executed applications and can be used to extract potential evidential data in case the tampering by the CEO has led to the registry values getting overwritten.

Carved files can provide additional information regarding files that were deleted by the suspect but get recovered during the process of the forensics examination as only the file metadata had gotten deleted and the actual contents of the file were still present on the disk. Orphan files, on the other hand, are files that have their metadata left but the actual file contents have undergone deletion. Both of these varieties of data recovery or metadata recovery help in the recovery of potentially crucial evidence related information that the suspects wanted to get rid of to prevent themselves from getting caught and charged for the same.

Analysing memory artefacts is only possible if the suspects' computer is found in a turned-on state and does not get turned off. It could lead to the discovery of potentially crucial data stored in memory, such as recently used passwords, relevant hashes, running processes, etc. The document related files and downloaded emails that are present on the system should undergo thorough scanning for potentially suspicious activity that links the suspects to the case or provides a brief overview or brings to light the issues related to the topic of accusation.

For performing most of these tasks, Autopsy or FTK toolkit serves as a good standard, primarily when used along with tools like RegRipper and other utilities from Nirsoft which aid in the easy discovery of most of the data mentioned above. For performing memory forensics, however, other toolkits such as the Volatility Framework needs to be used. Scanning of documents and emails for possible suspicious activity needs to be performed on an e-discovery platform to ease the task rather than making it tiring, cumbersome and erroneous. Such platforms are Nuix, Relativity, etc. which can aid in large scale document and email related data processing apart from adding connections between subject matters, therefore, making it easier to deal with large volumes of textual or document related.

3. **The work email accounts of each suspect** – The data present here should get added to the e-discovery platform mentioned earlier, and the data should get processed along with the other textual content so that the same is easy to deal with and automatic relations can be built by the software to further aid in the e-discovery process. The e-discovery process undergoes further explanation in **Answer 2 (A)** section of this document.

4. **The central server** – Since this is a Windows server, the same methodology can be applied as applied in the Desktops section apart from also checking the internal logs of the server that were present, and the databases and backups present for inconsistent entries.
5. **The email server which has Exchange installed** – This is the same as the central server and runs on a Windows platform apart from having Exchange installed. A few Exchange server-specific forensics practices need to get followed. These include extraction of logs from the Exchange server to display incoming and outgoing mails along with their respective timestamps. This forensic practices for the same also include combining the current and the backup data to gain extensive insight over the potential evidence present, which may have been tampered with, etc.
6. **The printers present in the office** – Printers generally do not contain much relevant information, but they contain the list of the last few printed documents in their cache memory along with the documents themselves. This method of document recovery is highly useful if the documents have gotten deleted or have been password protected. Additionally, if the printer is advanced enough, maybe additional details including the last user that used the printer, and a lot more of the documents recently printed can get acquired from it.
7. **The network connecting multiple devices** – It is essential to perform analysis of the network connecting the multiple devices as the routers could have undergone modification, and the connectivity may thus have gotten disrupted or restricted. This modification can easily get performed by the introduction of custom available firmware over the internet and then flashing it to the network components. Additionally, the switches connecting the internal network could potentially contain some messages in their message buffer or cache memory which could get used in identifying specific messages shared by the suspects or events and activities performed and the time of these particular network activities, primarily if they are related to the internal central server.
8. **The printer server storing print logs** – The printer logs are minimal as they only hold the data regarding the last few printed documents, however, a server coordinating with the printer logs every bit of data regarding every incoming document to the printer that gets printed. Additionally, a copy of the document in the image format (the format that the data is getting sent to the printer in the raw state) may also get saved providing additional insight into potentially deleted or password-protected documents. Apart from this, the general desktop-based forensics mentioned above can also get performed for uncovering additional important information, however, in this case, it might be useless as no one had hands-on access to the printer's server, unlike the central server. Thus, it might not have undergone tampering or would not contain sensitive information, thereby producing uninteresting results.

9. **USB sticks and DVDs found** – The USB sticks and the DVDs if re-writable need to get inserted into write blockers to take a forensic image on which the analysis can get performed. These devices may contain crucial information and thus are essential to the case. In this case, multiple documents have undergone discovery on the USB device. These documents should get added to the list of documents found on the desktop and the emails, to the e-discovery platform where a further connection could get made amongst the documents and their contents which could result in uncovering potential evidential data from a large number of textual documents.
10. **The mobile phone found** – The mobile phone, in this case, could play a massive role in providing additional information regarding the case. Firstly, a set of call logs could be requested from the network service provider. After that, a forensic image of the mobile phone could get taken so that further examination can occur. Next, the call logs could get tallied with the actual records of call logs presented by the network provider. This process would result in the calls that the suspect wanted to hide from the higher authorities. One can then cross verify the phone numbers with the phone numbers of the clients for additional confirmation towards the suspected scenario. It is also vital to perform forensics on the mobile device, which is a much more interesting albeit complicated task to generate a timeline of activities performed on the mobile phone. One should also gather additional data from the gallery, WhatsApp or other such commonly used applications to link the conclusive evidence found there and present it in the case if there is a need for the same. The web history, downloads, documents present and other such information, should get thoroughly checked in case of a smartphone for additional pointers which strongly link the accused to the case.

The password protection can get removed from the Excel, and Word files present by automating the process by uploading the files to a third party Word or Excel password cracking website and paying a certain sum of money or by using in house tools to brute-force the password so that effective e-discovery can take place. To reduce the time involved in the password cracking, the methodology used to open password protected files would be understanding the parameters involved in the same. This can easily be identified by converting any password protected excel or word file to a ZIP file and extracting its contents to view the XML information present in the EncryptionInfo file present within it. The process is depicted in the figure below.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<encryption xmlns="http://schemas.microsoft.com/office/2006/encryption" xmlns:p="http://schemas.microsoft.com/office/2006/keyEncryptor/password" xmlns:c="http://schemas.microsoft.com/office/2006/keyEncryptor/certificate">
  <keydata saltSize="16" blockSize="16" keyBits="256" hashSize="64" cipherAlgorithm="AES" cipherChaining="ChainingModeCBC" hashAlgorithm="SHA512" saltValue="qSVteebv18zGPqLjwEQ==" />
  <dataintegrity encryptedMacKey="uR0nJL3MBPmW0VBHL4p/xrh6TUHzmqY8H0CzadewDRebqVLQZKvWpuVaqC0XKICIEZHvGaUzeCLBKctFWtg==" encryptedMacValue="Br2nholM6G9WkfSkordIn8ZhsP9nokVh5pTnvUY6/Xg4qTn8x44cSH80V7Zj48yarBUs1J2NHNz170wojzQ==" />
  <keyencryptors>
    <keyEncryptor uri="http://schemas.microsoft.com/office/2006/keyEncryptor/password">
      <p:encryptedKey spinCount="100000" saltSize="16" blockSize="16" keyBits="256" hashSize="64" cipherAlgorithm="AES" cipherChaining="ChainingModeCBC" hashAlgorithm="SHA512" saltValue="QRA186wPkc9ttJhbkpOQ=="
        encryptedVerifierHashInput="Om47mFRm1VCdH4cdRqeg==" encryptedVerifierHashValue="m9FiyNmU8ukd2u3gall5v8NkclKqpc4m0HC6+9v8m7Z18MUU2F84mN3c8GvHvSP4625XZoSRL1XC8g75K4mY==" encryptedKeyValue="tm3M33EtcjB0KL+N5QPaEC5dHb8U5+87QW9y6y/ho=" />
    </keyEncryptor>
  </keyEncryptors>
</encryption>
```

**Fig 1. EncryptionInfo XML Data**

Each of the fields specified within this XML entry have a unique usage and the usage of the same is presented on this link ([https://docs.microsoft.com/en-us/openspecs/office\\_file\\_formats/ms-](https://docs.microsoft.com/en-us/openspecs/office_file_formats/ms-)

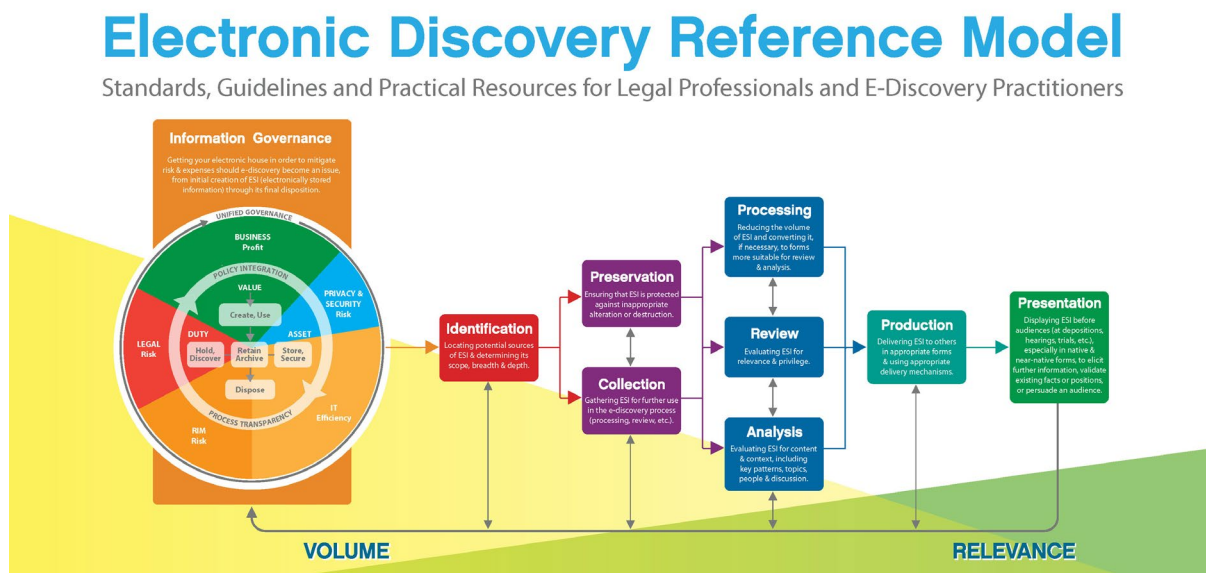
[offcrypto/87020a34-e73f-4139-99bc-bbdf6cf6fa55](https://offcrypto.com/87020a34-e73f-4139-99bc-bbdf6cf6fa55)) which could get utilised in automating the brute force process in house while also reducing time rather than scripting a manual program that opens up Excel or Word to interface with its enter password textbox, which makes the process much slower.

The CEO of the company additionally wanted details regarding the Electronic Discovery Model, Technology-Assisted Review, and the concept of Expert Witness. The details related to all three of these get presented in the following sections.

## Answer 2 (A)

### The Electronic Discovery Reference Model

The Electronic Discovery Reference Model is a diagram that depicts the conceptual representation of the e-discovery process, which is iterative in its approach. The steps outlined in the process do not all need to get carried out during the process of an e-discovery objective, while some steps have to get carried out multiple times to fine-tune the results [1].



**Fig 2.** The Electronic Discovery Reference Model

What this model depicts briefly is that as the process of e-discovery transitions from the first stages to the final stage, the total volume of data to be analysed decreases while the volume of relevant data increases. The e-discovery reference model consists of nine stages, each of which gets described in detail below –

- 1. Information Governance** – The ring presented in the picture above depicts the information governance reference model from the e-discovery perspective. This stage deals with the implementation or assessment of data governance standards set in place to mitigate risks, liabilities and costs in a scenario where an e-discovery assessment is essential. The entirety of this process revolves around incorporating or reviewing the strategies set in place by various organisations to balance risks that information presents compared to the value that it provides. The stakeholders involved in this scenario includes three generalised categories of individuals consisting of business users, IT departments and legal, risk and regulatory departments all of whom need information at a variety of levels to operate the organisation. The outer ring defines the dependency of various sectors of an organisation on data for achieving unified governance. In contrast, the inner ring depicts a lifecycle that should get

followed to reduce risk and expense related to the data getting handled by enforcing methods of policy integration and process transparency.

2. **Identification** – The aim of the identification phase revolves around the development and execution of a plan devised by the legal team to identify and validate potential electronically stored information sources correctly. The first stage involves developing a strategy and plan for identification of key players and then implementing appropriate electronically stored information management protocols that are compliant with the rules laid down by the courts, state and federal legislatures, and the government regulators. The next step involves the establishment of the identification team, which is accountable and responsible for various aspects of the identification process. Next, the identification of relevant electronically stored information sources get identified which includes the identification of critical witnesses and custodians, determining the critical time frames, determining a list of keywords, identifying potentially relevant document and data types, understanding file storage and email system parameters, etc. Finally, a certification of the potentially relevant electronically stored information takes place where the verification of the same gets performed by the e-discovery team head or the counsel that is involved. Three main processes occur after the previous phases where the status and progress reporting deals with allowing management the ability to analyse projects on a case-by-case basis, the documentation for defensible audit trail deals with the production of the documentation and the demonstration of the fact that the identification process was defensible if it comes under question, and the QC/Validation deals with the verification that there were no loopholes in the entirety of the process.



Although represented as a linear workflow, moving from left to right, this process is often iterative. The feedback loops have been omitted from the diagram for graphic simplicity.

**Fig 3. Identification Process**

3. **Preservation** – The preservation phase deals with the preservation of relevant and essential data in ways that are legally defensible, proportionate, efficient, and auditable to mitigate risks. This process gets achieved following a series of steps starting from the development of a preservation strategy that determines what data needs to get preserved, how long the data should get preserved and how could the data be preserved in a legally binding way. Next, there should be a well laid out plan that suspends destruction of potential electronic data that could act as evidence to a case in the future which gets performed by devising a preservation plan. The preservation method deals with the type of implementation that an organisation undertakes to preserve potentially relevant data. Finally, once every aspect of the data preservation model is set in place, the execution of the plan begins. The background processes of status and progress reporting, documentation for a defensible audit trail and QC/Validation play a role even in this stage to ensure that the standards are getting maintained while there

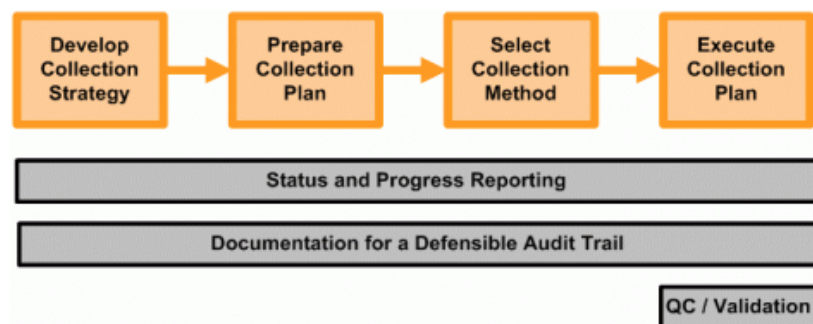


are no potential loopholes to the process so that the data could act as legal evidence in the near future if the need ever arises.



**Fig 4.** Preservation Process

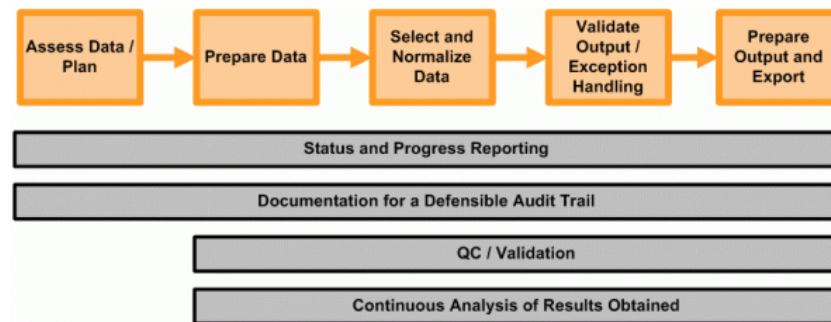
4. **Collection** – The collection phase deals with the collection and assorting of electronically stored information in legally binding matters of interest to the organisation, including government enquiries, litigation matters, etc. The process consists of the development of collection strategy, which defines how the data should get collected from the previously preserved sources or even otherwise from potentially relevant electronic information sources with the help of a collection plan. Next, the method for collection of data gets implemented or discussed based on the present stage of the organisation, so that the collected data is legally binding and the same can find its use in legal matters. Finally, the execution of the plan occurs as and when the need arises in the near future. The three background processes present here plays the same role as in the previous stages, and thus further explanation related to the same has been avoided.



**Fig 5.** Collection Process

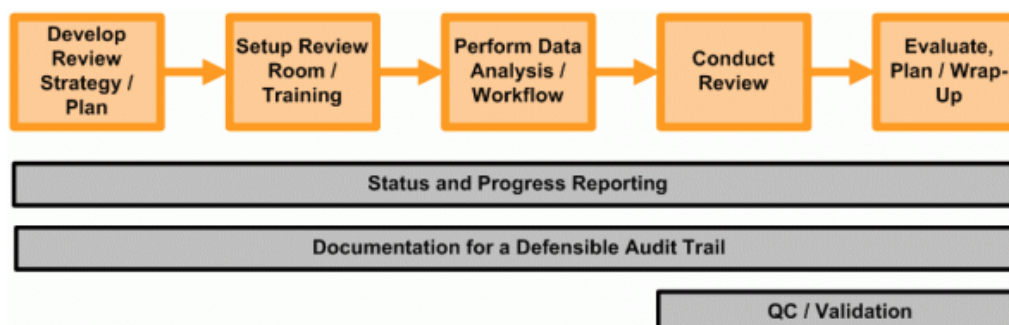
5. **Processing** – This phase deals with the identification and analysis of relevant electronically stored information for performing analysis and structuring based on the requirements mentioned. Multiple steps get performed in this stage, including metadata preservation, itemisation, normalisation of the format, and data reduction. The first step involves assessing the type of data and the plan set in place to handle such type of data. Once this gets performed, the data gets intermediately prepared for further steps such as the selection of the complete content or partial content of the data based on the rules set in place and its further normalisation. The normalisation process, in this case, deals with the bringing of data to a standard format equivalent to the other sources of data that undergo processing so that the same functions can get applied to each of them without incorporating additional measures to handle specific types of data. Finally, the results are validated, and exception handlers are set in place to deal with anomalies that occur as a result of the processing phase. Additional measures may get taken to process these varieties of

data if the need arises, else they get ignored. Once the data is completely processed, the same gets exported in a format on which global functions can get applied without individually crafting a function for a specific type of data. The background processes remain the same, playing the same roles as in the previous phases with the addition of the continuous analysis of results obtained phase, which deals with the analysis of the processed data generated and works on an iterative basis if the desired output does not get obtained thereby refining the data generated on each iteration.



**Fig 6. Processing Process**

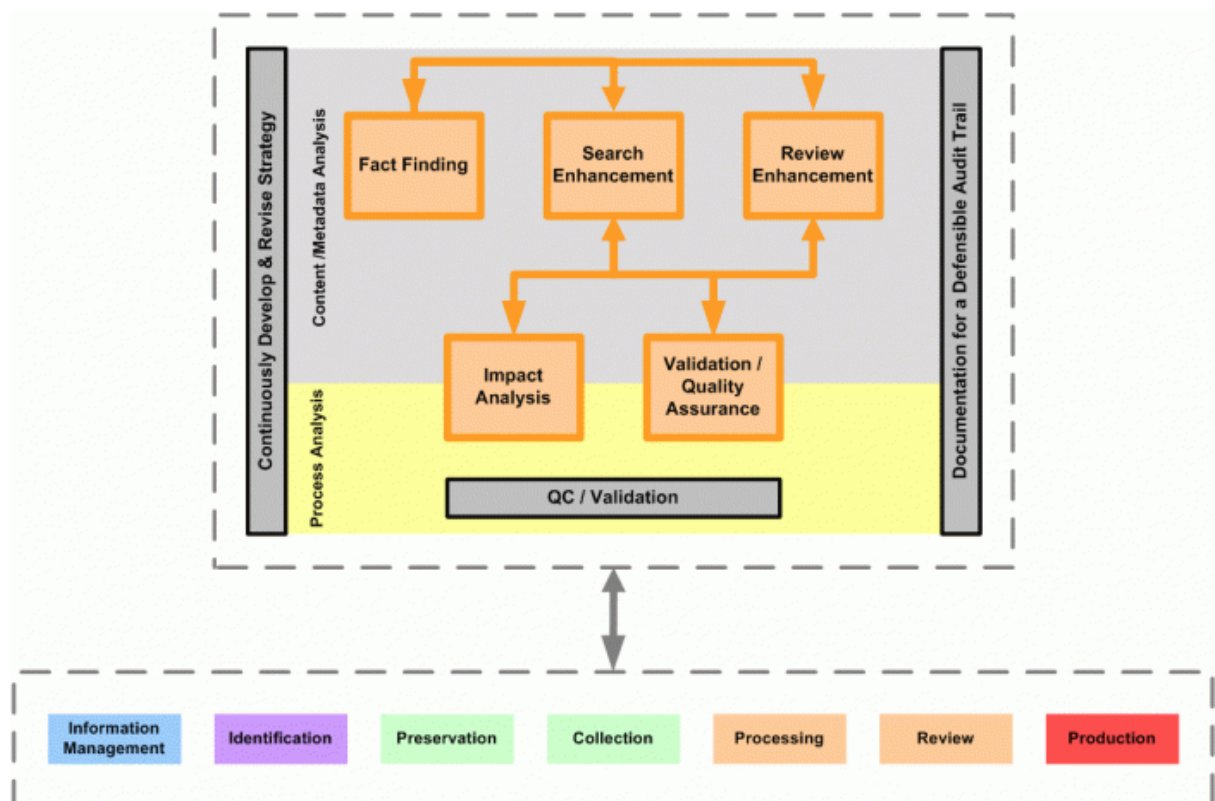
6. **Review** – The goal of this phase is to gain a deeper understanding of documents involved and categorising them into different subsections efficiently and cost-effectively. This phase gets performed to establish facts and links between the documents so that a more precise picture is visible to the legal team in the case of litigation. This phase revolves around the generation of a review plan or strategy that gets utilised for identification, categorisation and linking of multiple documents. The review room then gets set up where multiple instances of training take place discussing the requirements for the case and the categories of the division of data. Next, the data is electronically or manually analysed and categorised based on the plan previously set in place and the meetings held in the review room, where multiple individuals working towards the same case review the documents related to an e-discovery case to conclude the overall standing of the scenario and take further actions to assist in the process of the litigation. The background processes still exist in this stage for the same reasons as previously mentioned.



**Fig 7. Review Process**

7. **Analysis** – The analysis phase revolves around the validation of pieces of evidence gathered as a part of the previous processes, so that the legal team could take appropriate actions in the event of litigation. The other processes feed in data to this process which then gets utilised for actions such as fact-finding dealing with the identification of facts from the related documents. Other subprocesses play a significant role such as search enhancement which

deals with linking of multiple document contexts based on a similarity index so that when a search for a particular context gets performed, the rest of the related documents get linked in the process. The review enhancement deals with the updated review based on the litigations faced from time to time so that further detailed can be uncovered from the data present underneath. Impact analysis deals with the impact a particular document context has in respect to a particular scenario. The impact scoring gets done based on the impact that the reviewer feels a particular document has based on the analysis of the entire situation and the relevancy that the document plays in that specific case according to them. The validation and quality assurance assures the quality of the e-discovery in connection to a specific proceeding or litigation. The only change in the other background processes is the inclusion of the continuously develop and revise strategy which details the updates issued based on the reanalysis and rereview of the entire scenario as the case progresses.



**Fig 8. Analysis Process**

- 8. Production** – The production phase entails the production of the electronically stored information in a way as to reduce costs, risks and abide by the specifications provided within the mentioned timeline. The first step involves the identification of the form of production which must get produced. The next step is to perform the data analysis, which is essential to verify that the required result gets generated in terms of data. Next, the production requirements get analysed so that further preparation of files can get performed based on the necessary and relevant data. Finally, after the preparation of the relevant files, they are exported to the media device for usage in a litigation process. The same background processes are present in this phase as well, performing the same actions as in the previous stages.

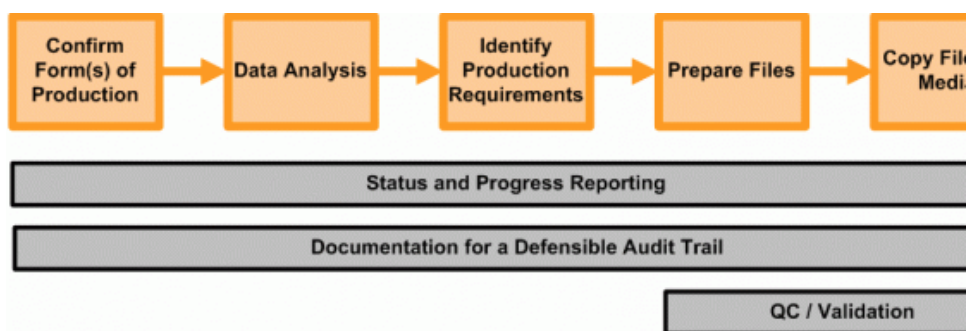


Fig 9. Production Process

9. **Presentation** – The presentation phase deals with the presentation of the relevant documents and the electronic evidence harvested from the sea of information for the process of trials and hearings. It consists of the development of a presentation plan or strategy that determines the most efficient and effective way of presenting a case depending on the scenario. Next, the format of the presentation of the digital evidence gathered gets decided, and then the exhibits are tested for validity and impact. If the exhibits get found to be useful, they then get presented as evidence, and irrespective of the entire process, all the exhibits get stored for further future enquiry.

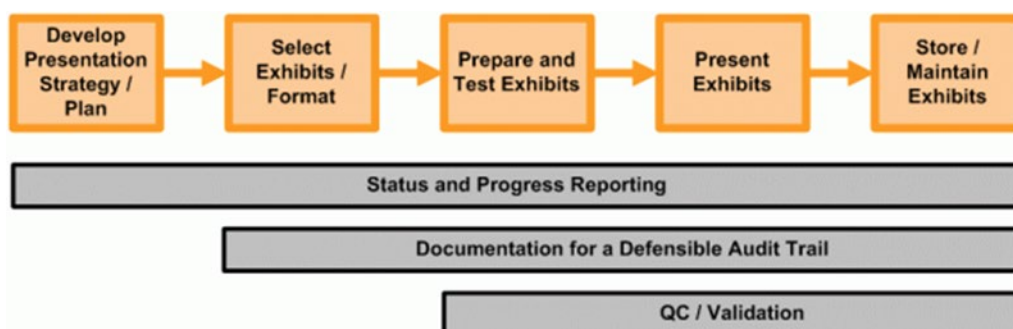


Fig 10. Presentation Process

### Technology-Assisted Review

Technology-Assisted Review (TAR) is a process of having computer software electronically classify documents based on input from expert reviewers, to expedite the organisation and prioritisation of the document collection. Multiple technologies present in the wild could undergo utilisation for technology-assisted review, but the top-notch competitors in the field are **Nuix**, **Encase eDiscovery**, and **Relativity**, respectively. All of these TAR tools have their ups and downs when it comes to data processing times, usability, training sessions, community, price-points, intended market, etc. A simple comparison of these tools in the graphical format based on reviews gets presented below with their explanations [2].




| Star Rating   |                   |
|---|-------------------|
|  Nuix eDiscovery Workstation | ★★★★★ 12 reviews  |
|  Relativity                  | ★★★★★ 143 reviews |
|  Encase eDiscovery           | ★★★★★ 21 reviews  |

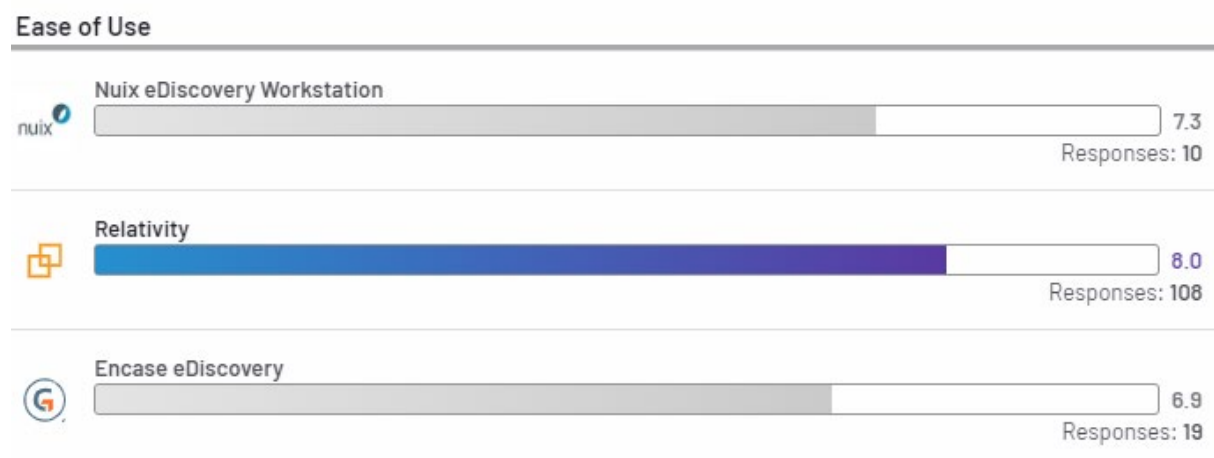
Fig 11. TAR Star Rating

Going by the star ratings, Nuix and Relativity both have received similar ratings from the community while the ratings related to the Encase eDiscovery is significantly lower. One additional point to keep in mind is that Relativity has the highest reviews, and thus, it is an excellent product based on the usage by the overall e-discovery community.



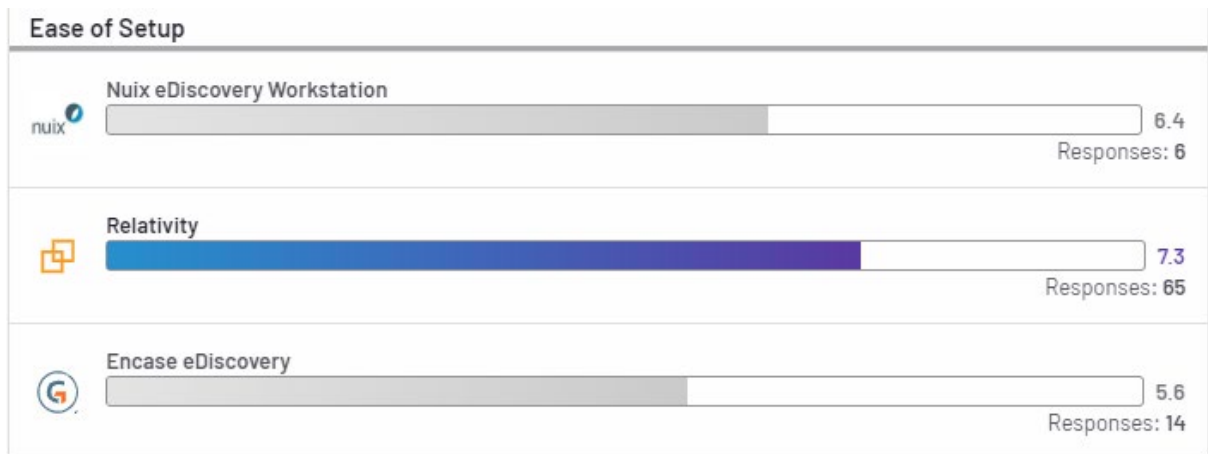
**Fig 12.** TAR Meets Requirement Metric

Comparison using this metric results in Relativity and Encase eDiscovery ending up in a draw on the scores, while Relativity has higher rating compared to Encase eDiscovery, while Nuix scores the lowest amongst the bunch.



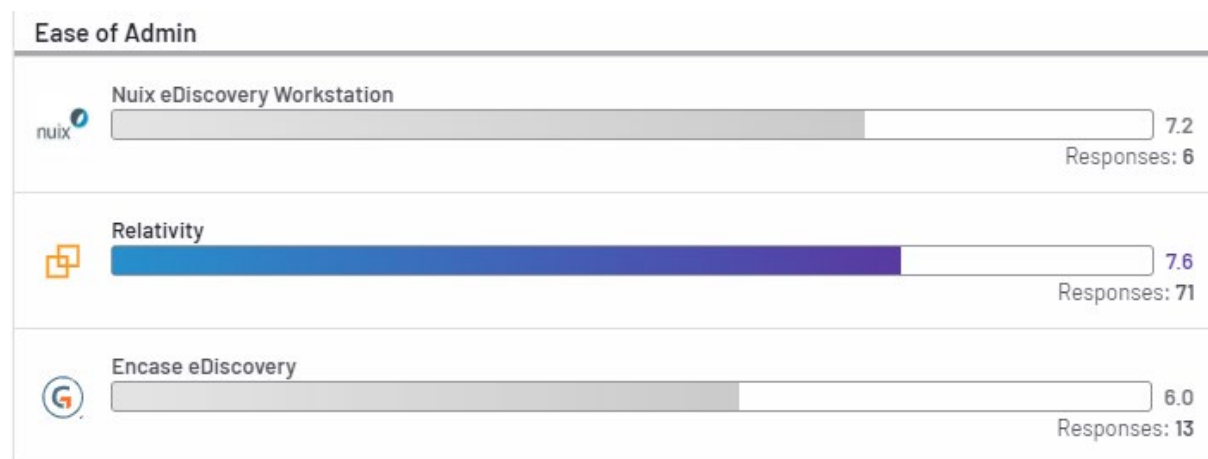
**Fig 13.** TAR Ease of Use Metric

In the ease of use metric, Relativity scores the highest with Nuix scoring the second position and Encase eDiscovery scoring the least suggesting that Encase is probably the least user friendly succeeded by Nuix and then Relativity.



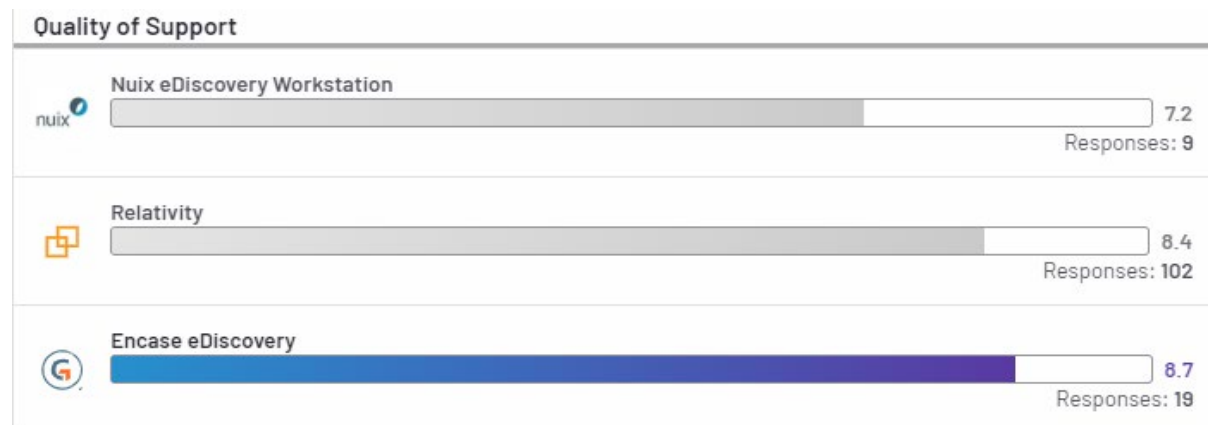
**Fig 14. TAR Ease of Setup Metric**

The ease of setup metric is topped by Relativity, followed by Nuix and finally Encase eDiscovery suggesting that the Encase product is the hardest to set up succeeded by Nuix and then Relativity.



**Fig 15. TAR Ease of Admin Metric**

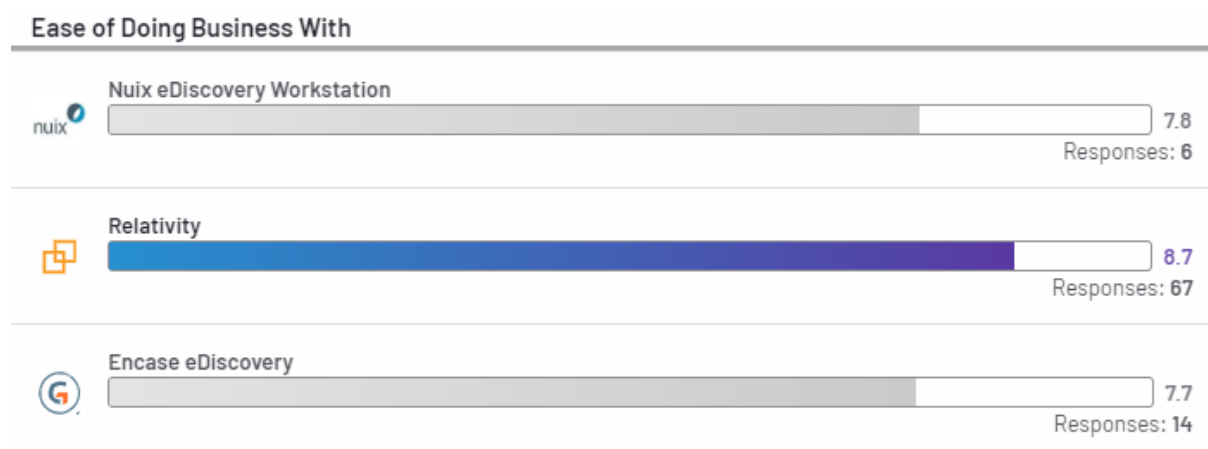
The ease of admin metric is topped by Relativity, followed by Nuix and finally Encase eDiscovery suggesting that the latter has the most difficult to use or the least features on the admin panel or login succeeded by Nuix and Relativity.



**Fig 16. TAR Quality of Support Metric**

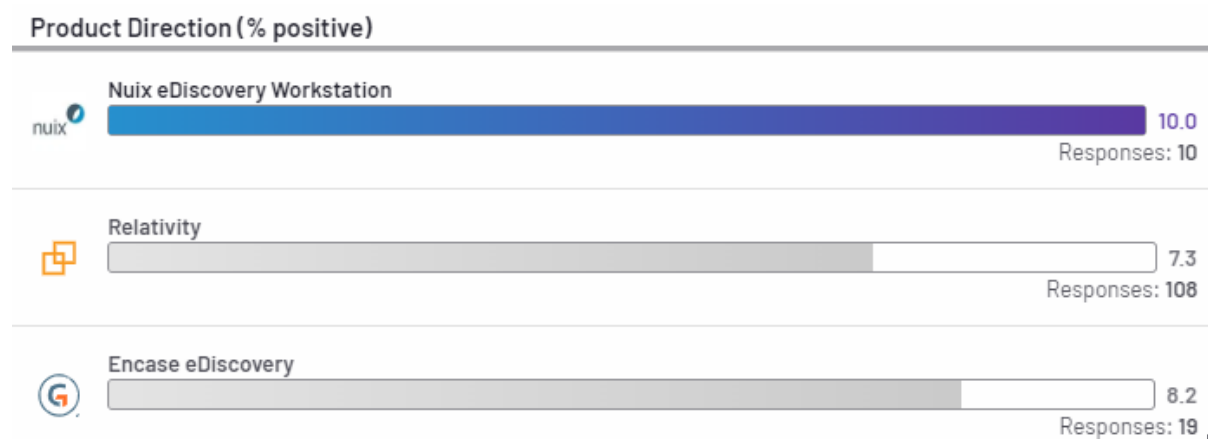


The quality of support metric gets topped by the Encase eDiscovery product followed by Relativity and Nuix respectively suggesting that Encase has the best product support amongst the TAR solutions, followed by Relativity and finally Nuix.



**Fig 17. TAR Ease of Doing Business Metric**

The ease of doing business metric gets topped by Relativity followed by Nuix and Encase eDiscovery respectively, suggesting that the business offered by Relativity is the best followed by the other two companies.



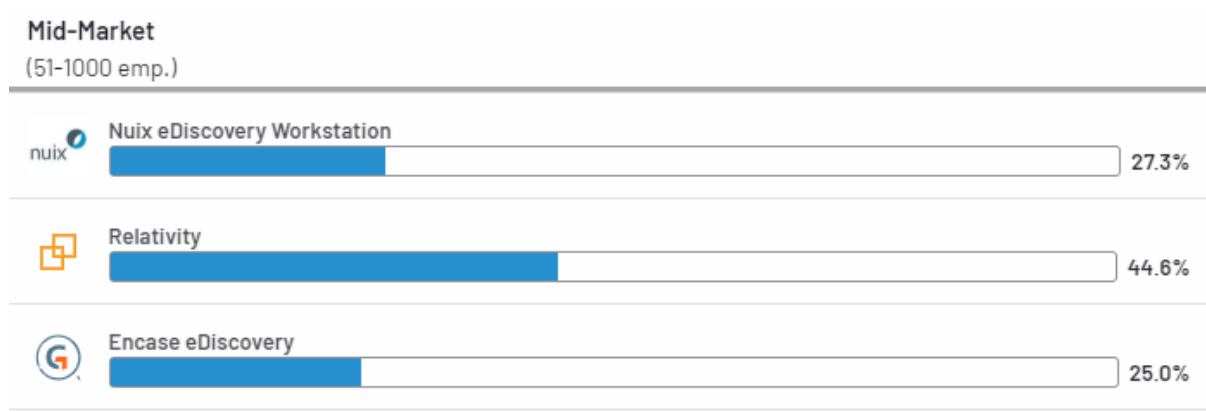
**Fig 18. TAR Product Direction Metric**

The product direction metric is topped by Nuix, followed by Encase eDiscovery and Relativity, respectively. Next, the usage of different products depending on the scale of the organisation is analysed.



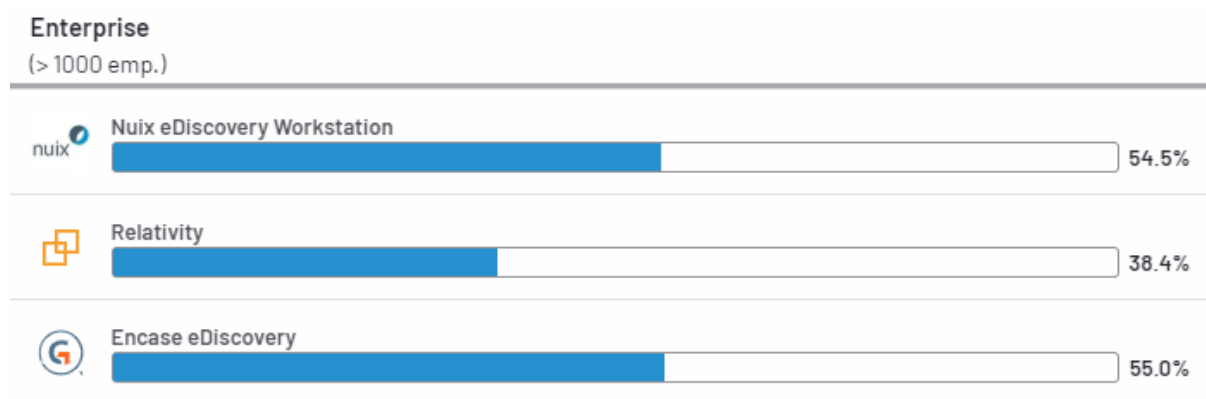
**Fig 19. TAR Small-Business Scenario**

As per the data gathered, Encase eDiscovery is the highest used product in small businesses, followed by Nuix and Relativity, respectively.



**Fig 20.** TAR Mid-Market-Business Scenario

When comparing the numbers of products utilised in the mid-market category, Relativity is the highest used product, followed by Nuix and Encase eDiscovery, respectively.



**Fig 21.** TAR Enterprise-Business Scenario

Finally, when considering enterprise scale of usage, Encase eDiscovery tops the list by a very slight advantage over Nuix, which is followed by Relativity.

Some additional advantages of these systems include –

1. **Nuix** – Boasts unparalleled speed, scalability and best in the industry data collection standards. It is also extensible by the use of multiple programming languages and gets trusted by organisations and industry professionals all over the world. Additionally, the company also holds multiple pieces of training to teach interested individuals.
2. **Encase eDiscovery** – The support of Relativity is the most significant advantage that this toolkit provides. Additionally, it supports other advanced features such as early predictability, advanced automation and unparalleled collections as detailed on the website, which is its selling points.
3. **Relativity** – The relativity community is probably the largest e-discovery platform community that exists with 180,000+ active users. They highlight features such as advanced security; cloud connects and open platform to attract customers on their website. Additionally, they also host training sessions for newbies in the field or to the software solution provided by them.



## References

[1]"EDRM Model", Edrm.net, 2020. [Online]. Available: <https://www.edrm.net/resources/frameworks-and-standards/edrm-model/>.

[2]"Pardon Our Interruption", G2.com, 2020. [Online]. Available: [https://www.g2.com/compare/nuix-ediscovery-workstation-vs-relativity?\\_\\_cf\\_chl\\_captcha\\_tk\\_\\_=7b0a2afdefa0ff6e4d02e44e5aa84fd56fd33901-1589916842-0-Ad851pZcAhtECNrSjQXqooskNiOCvE7-ApqNPgxYXgqq6wUHs1QeYPsOovUEUPV7yoY0C3bx4UWYDa3SGchNbQt\\_lIlg67ghh\\_jw29cBmZA07W2OA7vj11KF6PHMzl1biIIEq3YiVPe8RE5-2L3BfsRH5mDvRTPOIC\\_Op33hScRQZd2p8FfLxh1MZRCoZbFdmb5tJcCgs7JqJXaGMLuIY7f8zyGeIBKf\\_YbjQkkVLgktWHOZGFg5-P3s0KttJaXGxkYkyzs2JNEz6C\\_UMIPZ6F1692-Ut5BEics\\_JYbbezgM4VLuUcd-plcd38\\_zwSQmKJGPC8-2fSbMd4XMZFSaKgRipKZdew5CQDwO2b3cDju1U0jb\\_kemKcq-CC9ByohqdocPnWpXDWihHE8a5RE3yPPYb6PntPx7MHoEpWbSv7\\_\\_VH--QNgpT3u0CsqfA2AjoEeG\\_D9\\_2sSXzgBlT6pz6nhqAkYzOTtqdfjXCLTTszq3Q0ENfFHYrPJVD-ZHAX0ynZNTRWrNN0beBlzOUV-kQvXpMvAoO5HsIBcFXKDzIbgHf](https://www.g2.com/compare/nuix-ediscovery-workstation-vs-relativity?__cf_chl_captcha_tk__=7b0a2afdefa0ff6e4d02e44e5aa84fd56fd33901-1589916842-0-Ad851pZcAhtECNrSjQXqooskNiOCvE7-ApqNPgxYXgqq6wUHs1QeYPsOovUEUPV7yoY0C3bx4UWYDa3SGchNbQt_lIlg67ghh_jw29cBmZA07W2OA7vj11KF6PHMzl1biIIEq3YiVPe8RE5-2L3BfsRH5mDvRTPOIC_Op33hScRQZd2p8FfLxh1MZRCoZbFdmb5tJcCgs7JqJXaGMLuIY7f8zyGeIBKf_YbjQkkVLgktWHOZGFg5-P3s0KttJaXGxkYkyzs2JNEz6C_UMIPZ6F1692-Ut5BEics_JYbbezgM4VLuUcd-plcd38_zwSQmKJGPC8-2fSbMd4XMZFSaKgRipKZdew5CQDwO2b3cDju1U0jb_kemKcq-CC9ByohqdocPnWpXDWihHE8a5RE3yPPYb6PntPx7MHoEpWbSv7__VH--QNgpT3u0CsqfA2AjoEeG_D9_2sSXzgBlT6pz6nhqAkYzOTtqdfjXCLTTszq3Q0ENfFHYrPJVD-ZHAX0ynZNTRWrNN0beBlzOUV-kQvXpMvAoO5HsIBcFXKDzIbgHf).