5-2019

Analysis of The Typical Performance Routines for Recovering Data from Solid State Drive During a Forensic Acquisition

Bhargav Rajammagari
brajammagari@stcloudstate.edu

Analysis of the Typical Performance Routines for Recovering Data from Solid State Drive During a Forensic Acquisition

by

Bhargav Rajammagari

A Starred Paper

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Information Assurance

April, 2019

Starred Paper Committee:
Mark Schmidt, Chairperson
Dennis Guster
Sneh Kalia

Abstract

The increased adoption rates of Solid-State Drives as the primary storage devices in digital computing is causing traditional computer forensic examination to face enormous challenges. Digital forensics has typical routines and guidelines which are followed to extract critical data in the form of evidence from storage drives of digital devices like computers and laptops. These conventional mechanisms apply to the traditional spinning media disk-based hard drives as these devices leave artifacts through which the forensic expert can extract and recover the deleted files. However, the Solid-State Drives make use of NAND based flash memory implementation which makes it impossible to recover the deleted data from these devices. Solid State Drives are unique in nature and architecture which is steered by a controller chip inside the device. The controller and algorithms are kept as a secret by the vendors of these devices which makes it impossible to extract data from them. Over the years, the forensic department has found some routines and mechanisms to retrieve data from these devices. It is necessary to examine the challenges that these devices pose in data recovery and to examine the possible methods for data recovery from these devices.

Table of Contents

List of Tables

List of Figures

Figure Page

Figure Page

Chapter I: Introduction

Introduction

Digital Forensic is the art of retrieving data from digital devices like computer, laptops, mobile phones, for evidence as a part of criminal investigation. Digital forensic helps in solving criminal cases by extracting data from various digital devices and using the data as evidence in solving a crime investigation. Due to its validity in the investigation, many commercial organizations have used the support of digital forensic for their benefits in various types of cases such as bankruptcy, espionage, fraud investigations, and intellectual property theft. Research has shown that nearly 95% of the criminal cases leave evidence which can be captured and analyzed through the computer and digital devices (Bell, 2010).

However, with the advancement in computer technology, new forms of storage devices have emerged. So far, most computing devices use the spinning media-based storage devices where data retrieval was done using a forensic kit; even if the data got permanently deleted. With the advancement in technology, Solid State devices have come into use as the primary storage for most of the digital computing devices. Solid State Devices do not leave behind any artifacts inside them and pose a massive challenge for the forensic data team to recover the data from these devices. Over the years, the adaptation rates of Solid-State Devices have been increasing due to the extended data security that it provides, and this has been posing severe challenges for the forensic science department. This work examines the technology available for data recovery of these devices and makes recommendations about future technologies (Chen, 2011).

However, its usage has increased the complexity of forensic investigators. The TRIM, garbage collection and wear leveling creates sanitized disks which makes the data inaccessible for data recovery (Bednar & Katos, 2011).

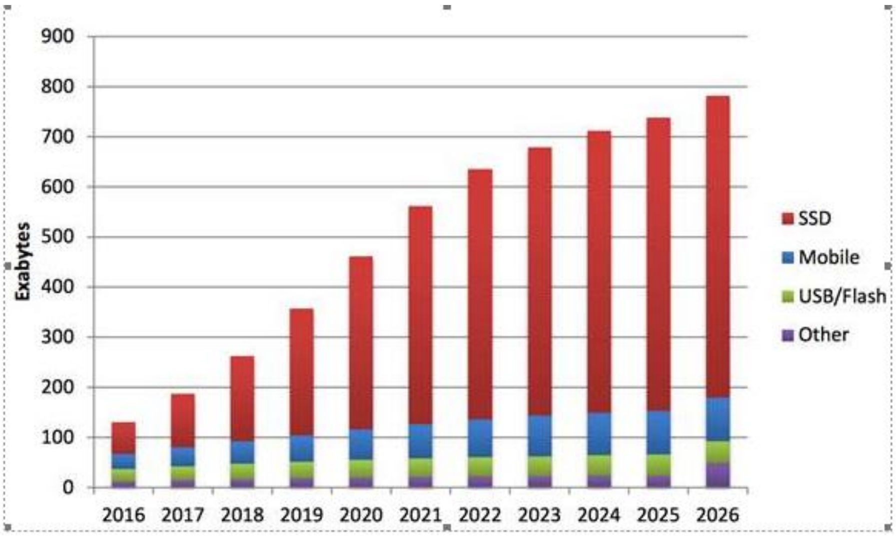

Figure 1. The Usage of Solid-State Drives (Geier, 2015)

Moreover, with each SSD, the garbage collection algorithm can vary in aggressiveness and can pose a severe complexity to the process of data recovery. It calls for an exploration of using various methods of data recovery that can be used by forensic experts on the solid-state drives. Based on these explorations, this work will examine the current state of the art of technology available for data recovery for these devices and will make a recommendation of future technologies (Bednar & Katos, 2011).

Problem Statement

Though SSDs are significantly fast in terms of speed compared to HDDs, Storage drives have brought revolution in terms of storage, which resulted in increased usage of this SSD. Storage drives raise a few challenges to digital forensic investigators, some of which are:

1. How can forensic analysis be done on SSD?
2. What routines are to be followed during a forensic analysis?
3. Which tools are used to perform the forensic analysis?
4. What type of data should we consider for this analysis?
5. How to compare the results between SSD and HDD?
6. How effective will the approaches be?

Nature and Significance of the Problem

The practices followed by digital forensic investigators to solve HDD data retrieval cases will be a lot easier and less challenging when compared SSD data retrieval because some SSDs are made using NAND flash memory.

As per the statistics shown in Figure1, we see a significant rise in the usage of SSD. So, the comparison of forensic analysis of HDD vs. SSD helps the researchers and investigators take better steps while doing a forensic analysis of SSD.

Objective of the Project

The research examines the complexity involved in SSD data recovery and the current state of the art of technology by analyzing the performance routines available in data recovery from these devices. The challenges forensic investigators face in extracting evidence during data recovery from the solid-state devices get easier when compared with hard disk drives.

Project Questions

The project questions are as follows:

1. What are the challenges faced by forensic investigators while using tools to retrieve information?

2. What type of files would be recoverable after deletion?
3. Would the deleted files be recoverable by other methods?
4. Can the results be the same with different approaches?

Limitations of the Project

The motive of this study is to examine challenges solid-state devices pose to the routines and mechanisms of forensic investigations. This study does not make any attempt to change any of the existing methods of extracting evidence or to the data recovery but suggests how these methods are adequate or not, for retrieving the data from these devices. The results obtained from this research might be valid only to this data.

Definition of Terms

Digital forensics: It is a process of interpreting and under covering electronic data. The primary goal of the digital forensics is performing a structured investigation by collecting, identifying and validating the digital information, so those past events are reconstructed while preserving any evidence in its most original form. Digital forensics can be very efficiently used in a court of law (Technopedia, n.d.).

HDD: A computer hard disk drive (HDD) is one of the mechanisms that steer the positioning, writing and reading of the disk. It is a non-volatile computer storage device which has magnetic disks or platters which are rotating at high speed. It acts as a secondary storage device. All the data stored non-volatilely are retained when the computer is turned off. Hard disk drives are also known as hard drives (Technopedia, n.d.).

SSD: Solid state drive (SSD) is an electronic storage drive built on solid-state architecture. SSD shares the same purpose as of Hard Disk Drive (HDD). NAND and NOR flash memory are used to make SSDs and to store non-volatile data. SSD is also called as an electronic disk drive or Solid-State Drive (Technopedia, n.d.).

Summary

In this chapter, we learned what HDD and SSD are and how they function, and how the market for SSDs is rapidly rising. We have also discussed the main problems faced by forensic experts for data retrieval in these devices. A brief description of the definition of terms also has been considered along with a short intro to study questions.

Chapter II: Background and Literature Review

Introduction

All the crimes related to computer technology and storage devices need digital forensic expertise to solve the crime investigation. A significant component of these storage devices is SSDs and HDDs. Retrieval of lost data from these storage devices is the primary job of a digital forensic investigator, especially when data retrieval from SSDs has become a difficult task. In this chapter, we will discuss why the data recovery process from SSD is different from that of HDD, and about the use of forensic tools with an in-depth functionality of HDD and SSD.

Background Related to Problem

Since SSDs have special features like TRIM function and wear leveling, and since they are made up of NAND flash memory, the techniques to study these behaviors and methods to extract data are very scarcely known. These features could be of significant advantage in fast in reading, write and store operations of SSDs, but possess a severe threat in the retrieval of deleted information.

Literature Related to Problem

Forensics. Digital forensics is a combination of elements in law with computer science and technology to analyze and collect data from computers, networks, wireless communications, and storage devices in such way that is acceptable as evidence in a court of law (Ries & Hill, 2017).

The field of digital forensics involves the exploration of digital data from digital devices as evidence for any case. The digital forensic thus collects and analyzes the data from computing devices like mobile, computer, laptops for obtaining evidence for a legal matter. For the past 30 years, this method has helped in solving various kinds of investigations for the public and the corporations as well. Digital forensics also includes the discovery of new information from digital computing devices which handle sensitive data. It also provides measures to ensure the data integrity, so that the confidential data cannot be corrupted by others and is intact for the evidence purpose (INFOSEC, n.d.).

It is through the digital forensics that many cases of fraud, theft, software privacy, software hacking, cybercrime, blackmailing, terrorism, prostitution, child pornography, domestic violence, trade secrets have investigated worldwide. Through digital forensics, cases of network intrusion, hacking, password hacking, cyber warfare, online trading frauds can also be examined from the computation device history and memory devices and can be used to recommend creating preventive applications to act as a safeguard against such crimes (INFOSEC, n.d.).

A few of the tools used in a forensic investigation are: (a) SANS SIFT, (b) Pro discover Forensic, (c), FTK, and (d) Autopsy (INFOSEC, n.d.).

Spinning media drives and file storage (HDD). Information in the computing devices is stored in the form of patterns of series of 1s and 0s. The 1s and 0s are used only for human interpretation but do not exist. The 1s and 0s signify the presence or the absence of electric charge or magnetism in the physical device. In today’s computing devices, the main memory, which is often called as the Random-Access Memory stores the patterns of 1s and 0s by conducting charges in tiny battery like capacitors. To write the data into memory, some of these memory cells are identified and selected. These cells are filled with electric charge as per the data binary code. To read back the data from memory, the memory controller reads the electrical charge present in the cell, and the application program generates the binary value for the charge stored in the memory cell. However, this storage of data in the form of charges poses a few issues. The primary problem being that the charge cannot be stored permanently, for an extended period. Hence it is necessary to read out the data and to recharge the memory cells to ensure that the data can be preserved in these drives. Such memory is called volatile memory. It makes use of a circuit to charge the memory cells and ensure that the data is stored in the device for an extended period. However, there are some non-volatile memories which can store the data for a permanent time without the need of the refreshing circuitry. These devices are called secondary storage devices, and the popular ones are the spinning media-based devices, i.e., the hard disk drive (Woodford, 2018).

Figure 2. Hard Drive (Phelps, 2012)

Previously commercial computers made use of spinning magnetic drum-based memories, which provide access to the data in the form of blocks of digital codes. These digital code blocks were written to and read from them randomly. It meant that the reading and writing from these devices in any order was helpful for the computer programs for smooth execution and eliminated the need for data to be read from any external storage devices from the beginning until the needed block found inside it. It also allowed changing the data on these storage devices through the process of re-writing the tracks of media. These Magnetic drums made use of the random organization of the data, and this helped in mapping the data to individual records of codes on the drum with clear information. Each of these tracks further divided into sectors (Woodford, 2018).

Slowly, this gave the concept of hard drives on the computer. Hard drives contain a shiny plate which is circular and has magnetic properties. This plate is known as the platter, and this is divided into small units of areas. These areas of platters can get magnetized independently and store the digital value equivalent to one or demagnetized values representing a value 0. It is through the concept of magnetization, the information is stored in these devices, even when there is no supply of power (Woodford, 2018).

These spinning platter-based storage devices replaced the traditional spinning drum-based storage drives. Slowly these evolved into spinning disks of magnetic media which are formatted with multiple tacks instead of the parallel magnetic drums. The track subdivision was known as sectors, and these sectors hold the data. For each of these drives, there was a moving read head or a write head that allows each platter to read/write data from or to any sector. While doing so some time was spent to reach to the correct track and correct sector. The time spent is called as seek time, and the delay for finding the first sector of the cluster by the read or write head is known as rotational latency of the drives (Woodford, 2018).

Over the years, the data to be stored on these drives started to increase, and this called for storage devices with increased size and increased capacity of the drives. It resulted in an organizational scheme of the file-based system where the approach was to store the data on these sectors in the form of files. This file-based system made use of a file allocation table, that gave the identification of the type of the file, its address, its permission, its ownership. The file allocation table also contained an index to the first cluster where the data of the file is present on the disk. In this scheme, a small amount of storage space was needed at the beginning of each sector to store index points which would link to the next sector to be used by that file (Woodford, 2018).

The hardware and software controllers worked in synchrony to access a file and lookup the desired data file from the file allocation table. It then followed a linked list of sector data to retrieve data from the file and use it for a program. When the file was not needed anymore on the disk, one of the approaches would be to overwrite the file by finding each sector allocated by that file through the file allocation table. However, this approach is not used by the hard disk. A faster method that reduces the wear and tears on the disk is used for altering the file allocation table, which is to show the entire files in the recovery area, so other programs use that. However, it is a standard practice to leave the deleted files intact, and this is used by the forensic experts and their tools to recover the data (Woodford, 2018).

Mechanism of reading and writing in hard drives. The process of storing data on hard drives is not much of a significant problem as compared to finding the data on these devices. The hard drives make use of a proper storage mechanism to store data in the tracks. For storing any information on the drives, the computer considers the file allocation table, which tells it about the independent sector information. After locating a free sector, the read-write head is moved to the sector area across the platters to store the data. For reading any data, the address for reading is obtained by the program. The address is located using the file allocation program and the read-write head is moved to that address. The content from that sector is read, and the data is moved to data registers which were provided to the programs that need the data (Brendan, 2017).

Solid state devices. Solid State Devices are storage devices made up of NAND based flash memory chips where the data is read and written in the parallel form. The essential element in use in Solid State Devices is the NAND based flash memory which is semiconductor-based metal oxide memory. This memory is non-volatile and can provide enough density and speed for making this device to be a perfect device for the primary storage device (Bell, 2010).

SSDs makes use of a grid of cells which can send and receive any data. Pages separate the girds of the device, and it is in these pages where the data is stored. Pages are grouped to form blocks. In order to write data in SSDs, the page should be empty, and it avoids overwriting of the data (Wei, Grupp, Spada, & Swanson, 2011).

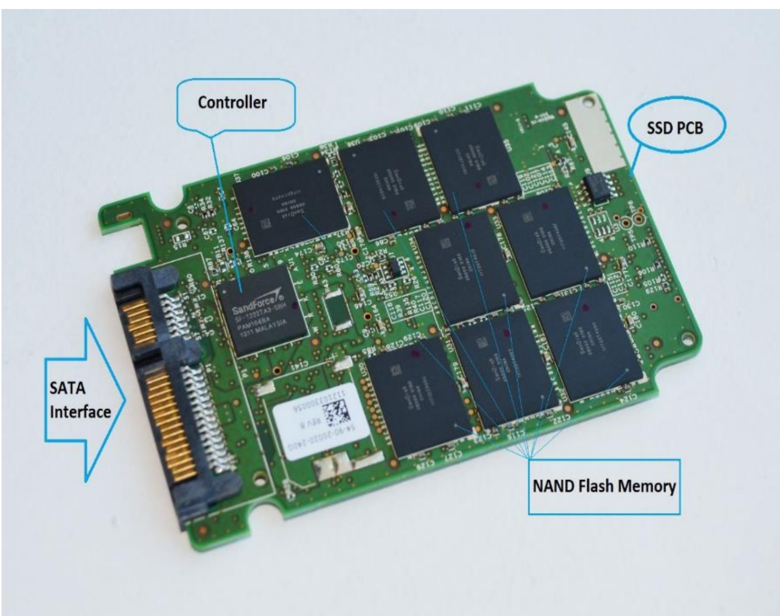

Figure 3. Solid-State Drive (C/net, n.d.)

The NAND flash chip components are stated in the figure. The flash page must be erased to make a write operation. However, the erasing mechanism takes a long time and can decrease the performance of the device (Wei et al., 2011).

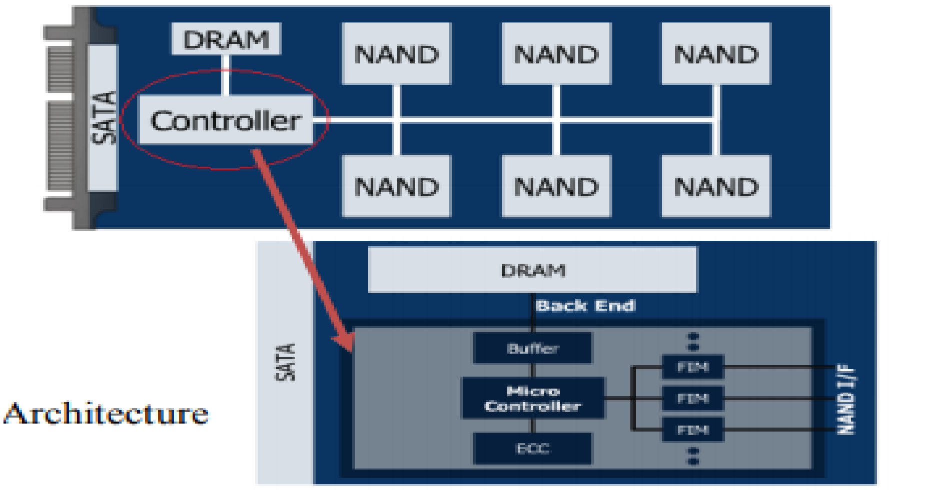

Figure 4. SSD NAND Flash Architecture (Thierolf & Uriarte, 2010)

The SSD central element the NAND flash poses several challenges to the SSD designers and manufacturers. Designers have increased the sensitivity of the read and write mechanism which gave rise to the Multi-Level Cell-based flash memory in SSD to increase the usage of solid state drives. Another issue was about to drive longevity due to the wear on the flash cells. This wear limited the erasing limits and the lifetime of the devices. It gave birth to the concept of wear leveling. It is through the wear leveling the block on the SSD is written once before any other writing to the block can be done again (Bux & Iliadis, 2010). Moreover, since the SSD makes use of the mechanism of reading data across all the chips at the same time, the need for a movable head has been eliminated. It results in no performance loss for the operation and wear leveling and has proven to be a valid concept for writing into the drives and increases its longevity (Chen, 2011).

SSDs makes use of a feature known as Flash Translation layer which runs in SSD controller that can translate the block commands to execute on the flash. This feature blocks the software access to physical blocks on SSD (Bux & Iliadis, 2010). The translation acts as a mask to the flash hardware on the operating system. Through this translation, the operating system may be reporting that the blocks of the SSD may be unchanged, but the fact may be that the SSD might be reallocating the data on the drives (King & Vidas, 2011). In the mechanism of the garbage collection, the collector moves the active pages out of the block and performs the erase function on these blocks. As the writes are being carried out on new pages, the garbage collector creates free pages where the write operation is performed. The usage of garbage collection and wear leveling has lead to the degradation of the performances as the garbage collector has to move between pages to keep the allocation pool filled. It gave birth to the TRIM function. TRIM changes disk garbage collection by allowing the operating system to mark blocks as deleted. TRIM changes disk garbage collection by allowing the operating system to mark blocks as deleted (King & Vidas, 2011).

SSD operation. The solid-state drives make use of flash memory like Random Access Memory (RAM), but the RAM is volatile and clears the data whenever the system power button is turned off. A solid-state drive, on the other hand, is non-volatile and retains the data even after the power loss (Chen, 2011). The solid-state drives make use of grid-based electrical cells for receiving data into the drives and sending data from the drives. Sections like pages separate these grids, and these pages hold the data of the drives. The pages are grouped to form blocks, and in solid-state drives, data is written into an empty page of the block. In Hard disk drives, the data can be written randomly to any location on the drive at any time signifying that data is overwritten, but in solid state drives data cannot be overwritten. For overwriting the data should be first written to some empty page and only then is that page used for writing the data. In solid state drives when there are enough unused pages, it will take the data of the block and commit that area of the memory to erase the whole block. Now the determined image will be printed on the block that has unused pages (Chen, 2011).

Working of SSD. The main two parts of the solid-state drives are its controller and the NAND flash memory. Along with other components, the whole circulatory for the solid-state drives is placed on a printed circuit board. The entire printed circuit board is placed in a casing and sold as solid-state drives (Chen, 2011).

The controller of the solid-state drives is a processor that is in the circuitry of the drives and act likes a bridge of the flash memory components and the host computer. The codes provided by the solid-state drive’s firmware are executed by the controller which contains information to fulfill the requests of the hosts. It is the controller that shall decide how the solid-state drives should function and what features it should offer. The main functions which the controller controls and executes for the solid-state drives are read function, write function, error checking function, deletion function, garbage collection function, wear-leveling function, encryption function, overprovisioning function and the RAISE function (Chen, 2011).

The next important component of the solid-state drives is the NAND flash memory. The NAND flash memory is an integrated circuit which can store information. Single layer NAND flash memory is used in enterprise-based solid-state drives while the commercial solid-state drives make use of multilayer cell NAND flash memories. The single-layered solid-state drives are faster in nature, and the last longer which makes it expensive than the multilayer cell NAND flash memories. Writing operation in the solid-state drives takes place when the controller programs the cells of the memory for storing the data. The memory cell holds the voltage in the form of 1 or 0 and stores the data in the binary form (Chen, 2011).

Reading process in solid-state drives is pure as the controller does not much to do in reading, but the writing process in the solid-state drives is a very complicated process. The NAND flash memory cells are programmed for a time, and after that it becomes unreliable. This property of solid-state drives is known as write endurance or program-erase cycle. Wear leveling is used to reduce the impact of write endurance on the drive, which ensures the effective use of chips cells that can be used one by one.

The solid-state drives do not provide the easy overwriting of the old data as in hard disk drives, and this inefficiency makes the data management in solid-state drives trickier than the hard disk drivers. Data is written in the form of pages by pages and is erased through a block by block. When the data is deleted in solid-state drives like by empty recycle operation no erasing takes place. Windows operating system makes use of a TRIM command which marks the data for erasing as an invalid page. However, the actual erasing of the data takes place only when the user writes some new data into the drive. There shall be no writing in the solid-state drives without erasing the existing data first unless the solid-state drives are a fresh piece. For the old drive, the process of wear-leveling and garbage collection is executed to enable the drive to write the data into it (Chen, 2011). Due to garbage collection and wear-leveling data is re-written from one place to another and this process is known as the write amplification.

Literature Review Related to the Methodology

TRIM. This functionality is used to erase blocks that are marked to be deleted by the operating systems. It has negative impacts on data persistence and forensic analysis. After deletion, the data cannot be guaranteed as the memory controller of the SSD decides the when and the number of marked blocks to delete. It is an essential function of SSDs which is not available in HDDs. One of ATA Set management command attribute is TRIM. The operating system using TRIM will inform the block that is to be deleted on the SSD. TRIM gives the list of blocks which are safe for removal from the device. In Windows Server 2008 or Windows 7, the trim command is enabled by default. The user can manually disable the TRIM command (Geier, 2015).

```
Enable:
fsutil behavior set disabledeletenotify 0

Disable:
fsutil behavior set disabledeletenotify 1

Check the status of TRIM:
fsutil behavior query disabledeletenotify
Results explained below:
DisableDeleteNotify = 1 (Windows TRIM commands are disabled)
DisableDeleteNotify = 0 (Windows TRIM commands are enabled)
```

Figure 5. Code Snippet for TRIM Functionality (Geier, 2015)

Garbage collection. Garbage collection works closely with TRIM functionality. It keeps track of the cells that are to be deleted and combines extra data of empty ones which helps in removing other cells. Garbage Collection works in the background and will only work with TRIM command. If the file system instructs the memory to delete the address or if the LBA address maps to a new address, the page is removed immediately, but instead, it is marked as deleted using TRIM command. The garbage collection keeps track of the files for removal. If all the files in the block are to be removed, the garbage collector erases the whole block. If most of the files in the block are to be removed, or if more empty blocks need to be created, the garbage collector moves the remaining items to different blocks and erases the complete block. This operation is performed in the background (Geier, 2015).

Wear leveling. Each NAND cell contains a limited lifespan and has a limited number of lifecycles and withstand more than 100000 cycles. Information is not always expected to get updated with the same frequency, some of which will take a longer time, and the other gets updated instantly. Maintaining the wearing out of cells is essential for the aging to be uniform and minimum. This process is called wear leveling. There are two approaches for wear leveling namely:

- Dynamic Leveling
- Static Leveling

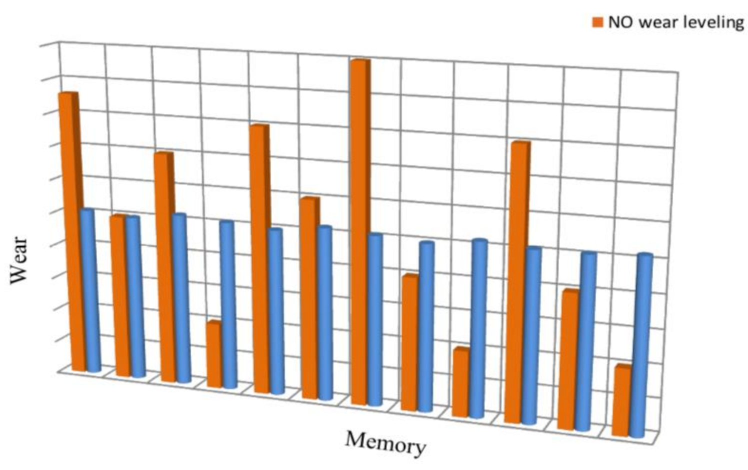

Figure 6. Wear Leveling (Geier, 2015)

Dynamic leveling is a technique of rampaging LBA address from one host system to the next available page. When the host updates the data, it either updates the page or writes into the drive. Data is always written into the next free cell with less aging level. Equal wearing is not guaranteed as unchanged cells will stay untouched (Geier, 2015).

Static leveling does the same functionality as dynamic leveling, but it also moves static pages to other pages periodically. Data in the oldest pages can be moved to an average aged page to make it usable for new data (Geier, 2015).

Forensic investigation of spinning media drives. Digital forensic experts rely on a few principles while examining the evidence from the spinning media devices. No harm should be done to the original media drive so that the integrity of the drive is preserved. To do so, the experts do not write anything on the drive but instead, make a backup copy of the drive and then perform the forensic analysis on the copy rather than on the original drive. Till date, a technique of dead box forensic was executed which considered the original drive as the Holy Grail and used other forensic tools to image the drive completely without altering its contents (Bednar & Katos, 2011).

Forensic investigation of SSD devices. The hardware of the solid-state devices, along with its firmware, exposes only those blocks of the device memory that are mapped to the sectors and attach it to the computer-based forensic applications. The forensic software examines the files which are allocated and are visible to the file system directory. Some of these files will be available in recycle bin. These are inactive files, and the fragments that are scheduled by the solid-state drive must be initialized to make them visible to the forensic analyzer software. It is found that the repeated usage of the forensic analysis software on the solid-state drives which is recently formatted may show different results while being executed multiple times as these fragments reinitialize themselves. On the other hand, the repeated analysis by the forensic analyzer application on spinning media based hard disk drives generate same results, and the operation of this software on these hard disk drives is reliable and stable contrary to solid-state disk platforms (Zhao et al., 2013).

The solid-state drives create complexity in the data recovery due to two main processes. First through the implementation of the wear –leveling function, the data on the solid-state drives are not written in logical, sequential blocks but are scattered on the blocks of the device. In case any block fails then the file data is written to a different location which is unallocated in use. It makes the sequencing and addressing of these blocks a complex issue. These sequencing and addressing of the solid-state drive blocks are under the control of the manufacturer's algorithm and creates complexity for the forensic expert to get the file sequencing and to address. The other complication arises from the re-initialization of the deleted files quickly since the hardware blocks the initialization of the "blocks" before writing. Due to re-initialization, the contents of the erased data blocks remain in the solid-state drive for a brief period creating a challenge for the forensic experts to recover the data from the drive (Bell, 2010).

In an experiment, it was found that in 30 minutes, nearly 99% of the deleted files from the solid-state drives were wiped off and were not available for the forensic recovery. The deleted files in the solid-state drives are available only for a short period and these, when rewritten, make the deleted files to be unrecoverable for forensic experts. In research, it was found that the TRIM command enables commands to eliminate the deleted files from the drive and makes their recovery impossible but in the absence of the TRIM command, some portion of the data can be recovered (Bell, 2010).

Another challenge that the solid-state drives pose to the forensic experts is the encryption of the data at the device levels. Many third-party vendors and manufacturing companies provide a service of encryption of the disk data to preserve the integrity of the data on the drive. Due to encryption, the controller of the solid-state drives provide another level of abstraction for the original solid-state drives data and offers a cluster-based view of the data to the host computer. This abstraction creates complexity in forensic data recovery (Bell, 2010).

Summary

The study of this chapter shows the importance of SSDs and hard drives. It also shows the exact functioning of both. The advantages and disadvantages of SSDs and hard drives and the process of information storage in both also have been discussed in this chapter. The importance of SSD in the forensic examinations and the main components of the SSD such as NAND flash memory and controller are given in this chapter.

The main components of a hard drive also have been discussed. It can also present the need and importance of both the devices in today's world. The solid-state devices are unique and architecture. A controller chip inside the device controls them. It is this controller and the algorithms which are kept as a secret by the vendors of these devices. These controller and algorithms make it impossible to extract data from these devices. This study examines the main challenges the forensic investigators face in extracting evidence or data recovery from the solid-state devices and analyses the typical performance routines for recovering data from Solid State Drive during a forensic acquisition.

Chapter III: Methodology

Introduction

In this chapter, we discuss the methodology that we use to retrieve the deleted data using digital forensic techniques. The tools that we use for this process also will be addressed, and the comparison of results between SSD and HDD will be provided. We also discuss the hardware and software requirements on the machines or devices that have been used to perform the analysis.

Design of the Study

A laptop with Windows operating system is required. The primary motive of this experiment is to retrieve the deleted files in SSD and HDD using forensic tools. They are formatted before we use them. Load data into both the devices and remove them. We use FTK Imager to create images of the disks respectively. The images are taken at regular intervals of time and analyze those images with FTK tool kit. Check how many files have been recovered in SSD and HDD, compare the results.

Data Collection

Data collection process involves a laptop with an SSD and an HDD. Some specific set of files have been taken as evidence and passed on to the disks HDD and SSD respectively. Both these evidence files have been deleted. Now the garbage data to each of the disks has been passed again and has been removed.

Continued sending of garbage data to both the disks on specific timelines have been done and have been deleted. Now the disk images of HDD and SSD have been taken at regular intervals of time and analysis of these images and checks on the number of evidence files have been retrieved in SSD and HDD is done, and the results are compared.

Hardware and Software Requirements

Operating system Requirements: Windows & and above, i.e., 10 version

Software Requirements:

- FTK Imager
- FTK tool kit
- HD Shredder
- Microsoft Office

Hardware Requirements:

- Laptop – Dell Inspiron 13 – I7-7500U
- SSD – 120 GB- Kingspec,
- HDD- 120 GB- Blueendless,

Timeline

Table 1

Project Timeline

Task	Duration
Selection of Topic by Reviewing Different Articles	**Three weeks**
Collecting Information Related to the Topic	**Two weeks**
Introduction and Literature Review	**Three weeks**
Review of the Document	**One week**
Project Proposal Documentation	**One week**
Collecting All the Resources for the Project	**Three weeks**
Creating Image Files and Analyzing the Image Files	**Four weeks**
Comparing and Analysis of Results	**Three weeks**
Documenting the Results	**Two weeks**
Project Defense Documentation and Presentation	**Two weeks**

Chapter IV: Data Presentation and Analysis

Introduction

Here we explain how the data has been collected by acquiring images of both external HDD and SSD. As a part of the next step, we will discuss on how we have extracted the data from the images acquired and how the comparison and analysis of the data have been performed along with the comparison of the results obtained from both the drives. The tools that have been used for this process are FTK imager and FTK.

Data Presentation

In this section, we would be discussing what type of files is present in the drives. There are different formats such as PDF, word documents, excel sheets. These have a high scope of evidence present in them as they might contain crucial information.

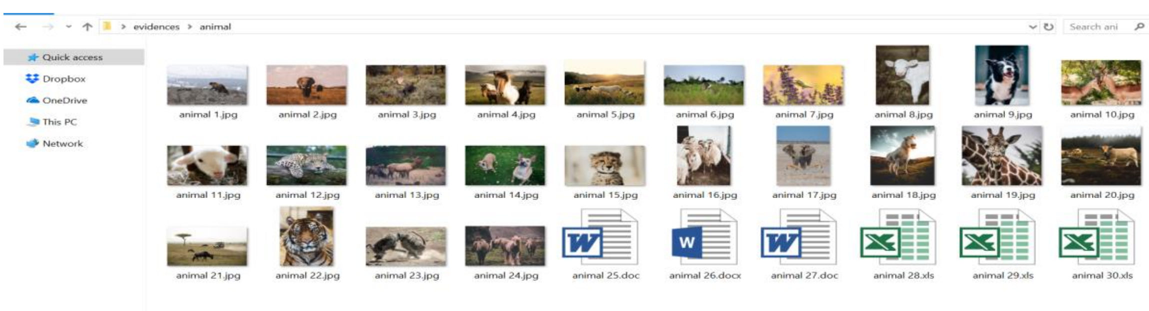

Figure 7. Contents of Animal Folder

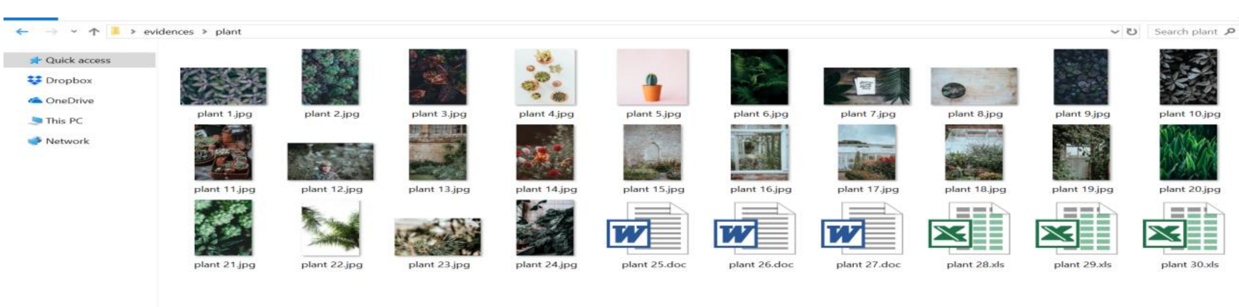

Figure 8. Contents of Plant Folder

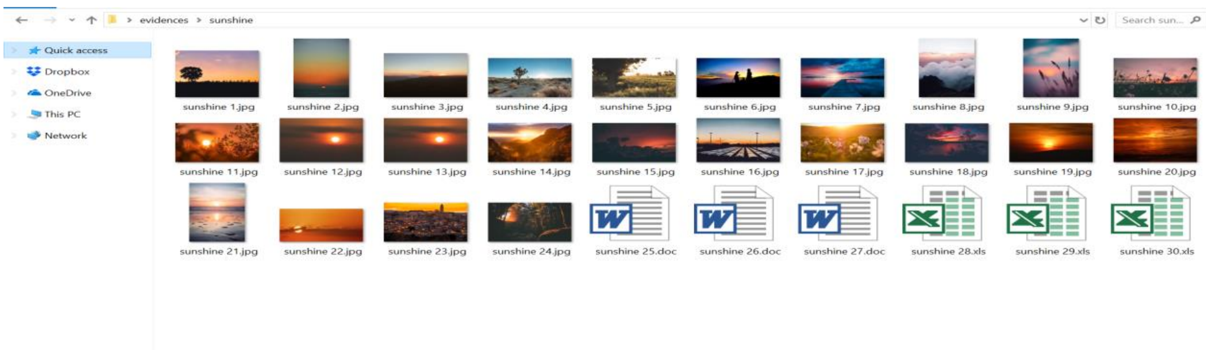

Figure 9. Contents of Sunshine Folder

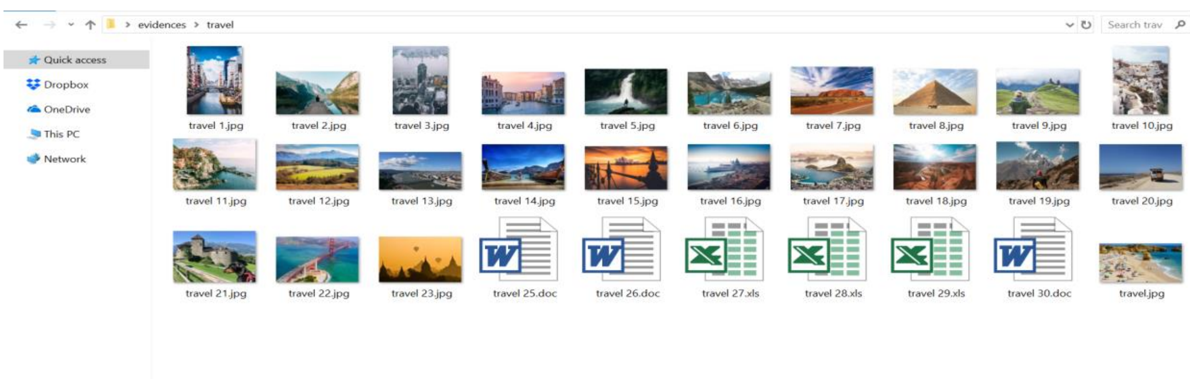

Figure 10. Contents of Travel Folder

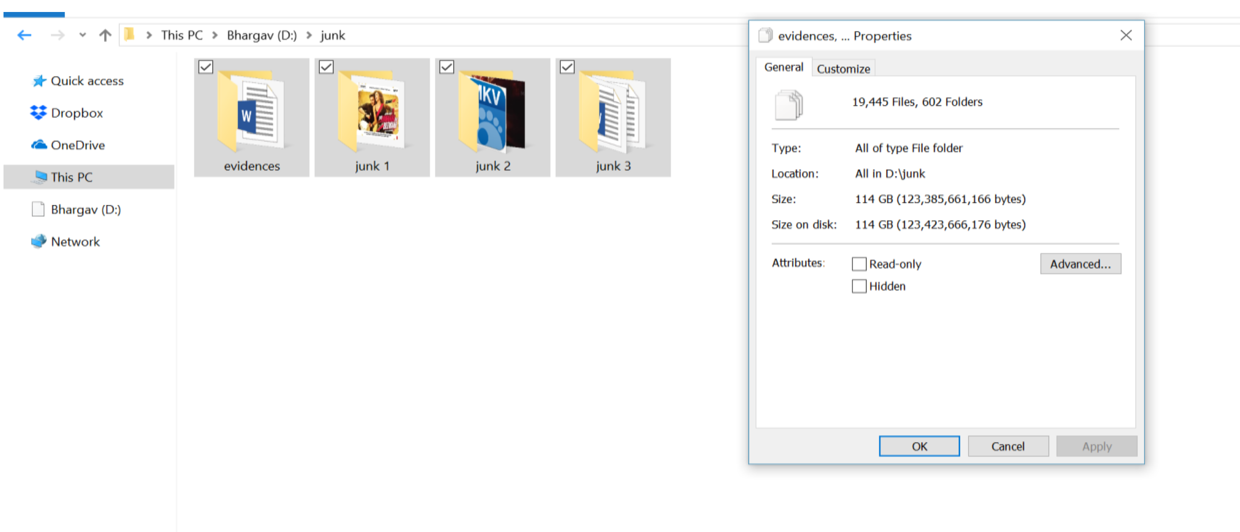

Figure 11. Evidence and Junk Folders

Evidence will be passed to HDD as well as SSD. The five Junk folders which have a wide variety of data like Excel sheets, PDF's, Images, MP3's with a size of 80GB will be copied to both.

First, we connect both SSD and HDD to the laptop, and we can see the preview as follows:

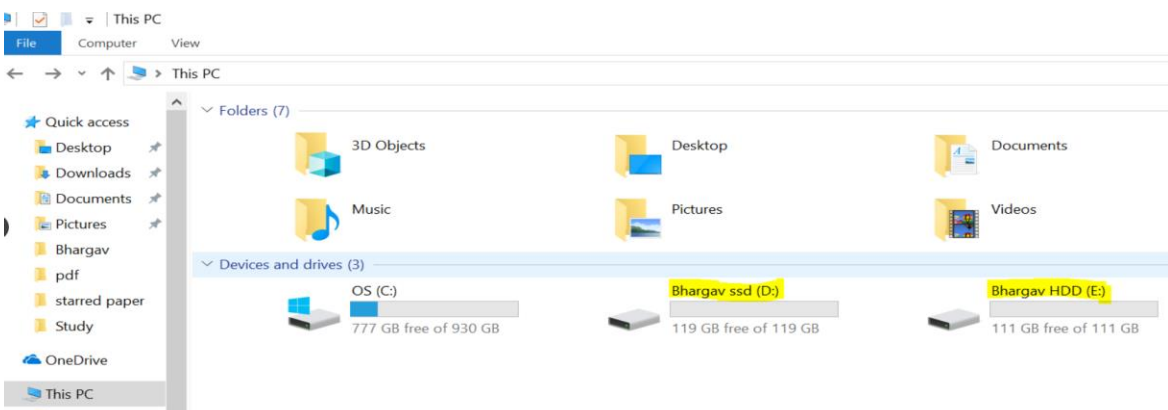

Figure 12. SSD and HDD Connected to a Computer

As a part of the next process, we copy all the evidence files and folders to both the HDD and SSD, and we can see the action below:

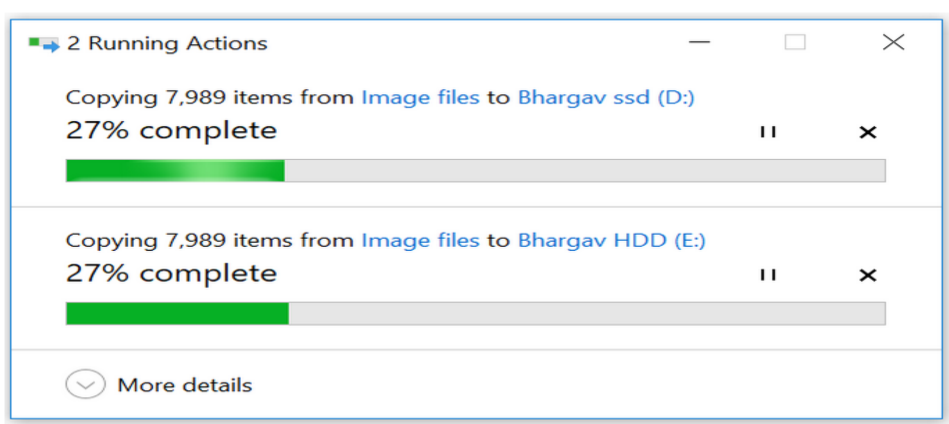

Figure 13. Copying Evidence Files to SSD and HDD

Evidence folder in SSD and HDD. We can see the evidence which is in the form of files and folders in the SSD

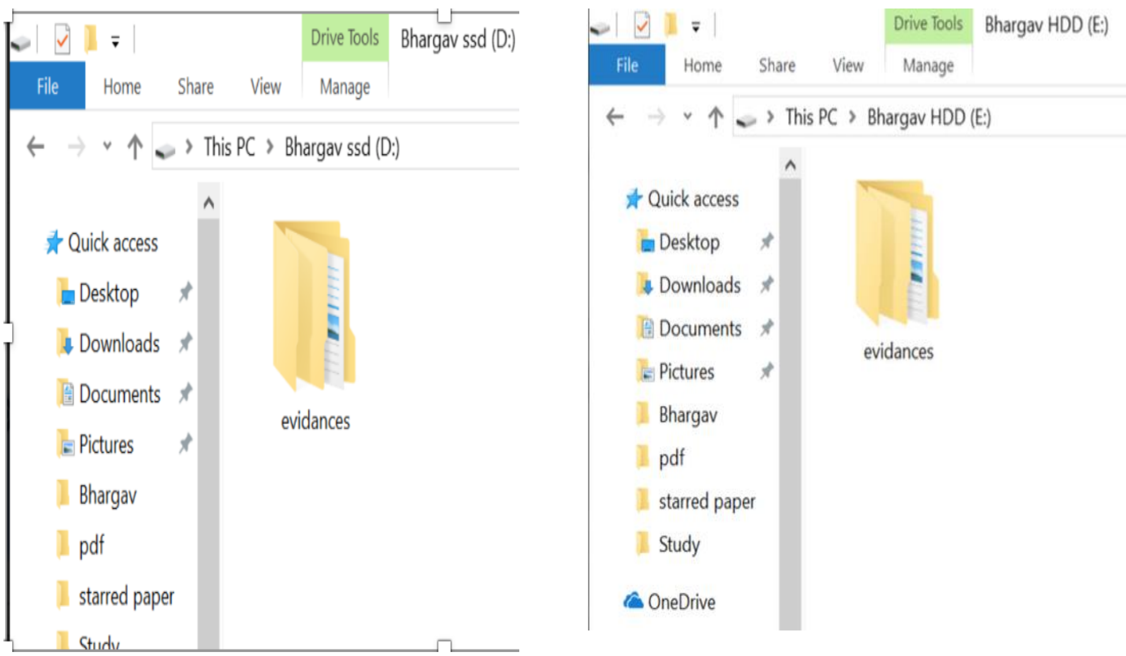

Figure 14. Contents of HDD and SSD

As there are Five Junk Folders, they are passed into SSD and HDD with different combinations as follows:

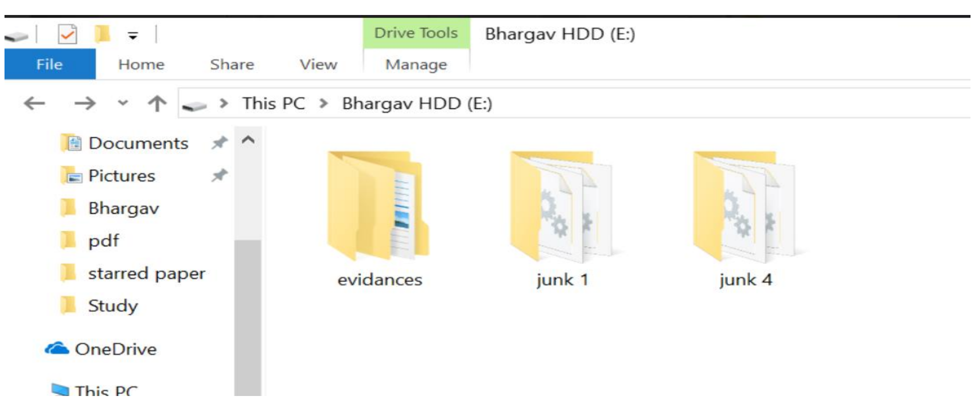

Figure 15. Contents of HDD

Delete the evidence and Junk folders and copying new Junk folders to HDD and SSD.

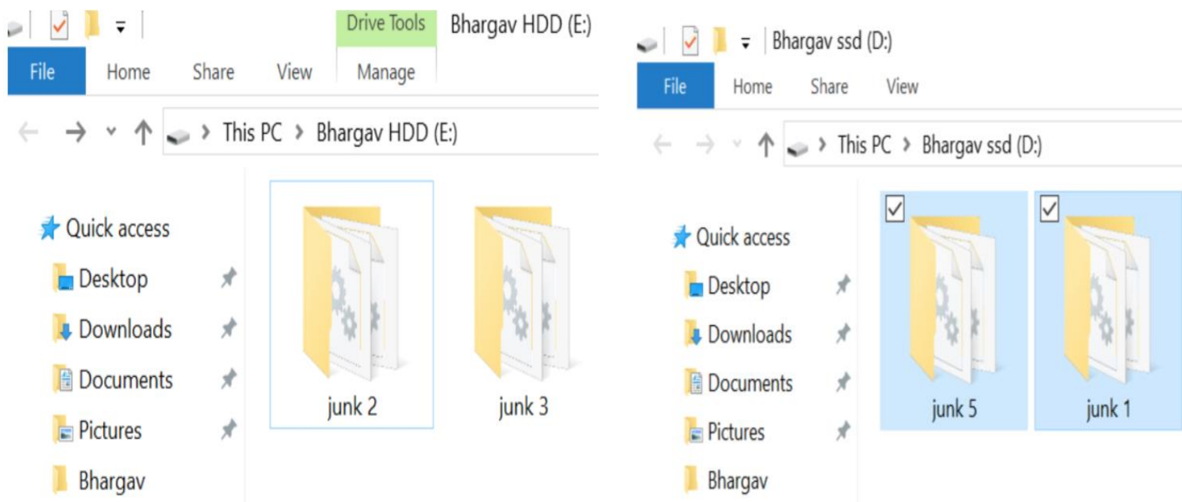

Figure 16. Junk Folders in HDD and SSD

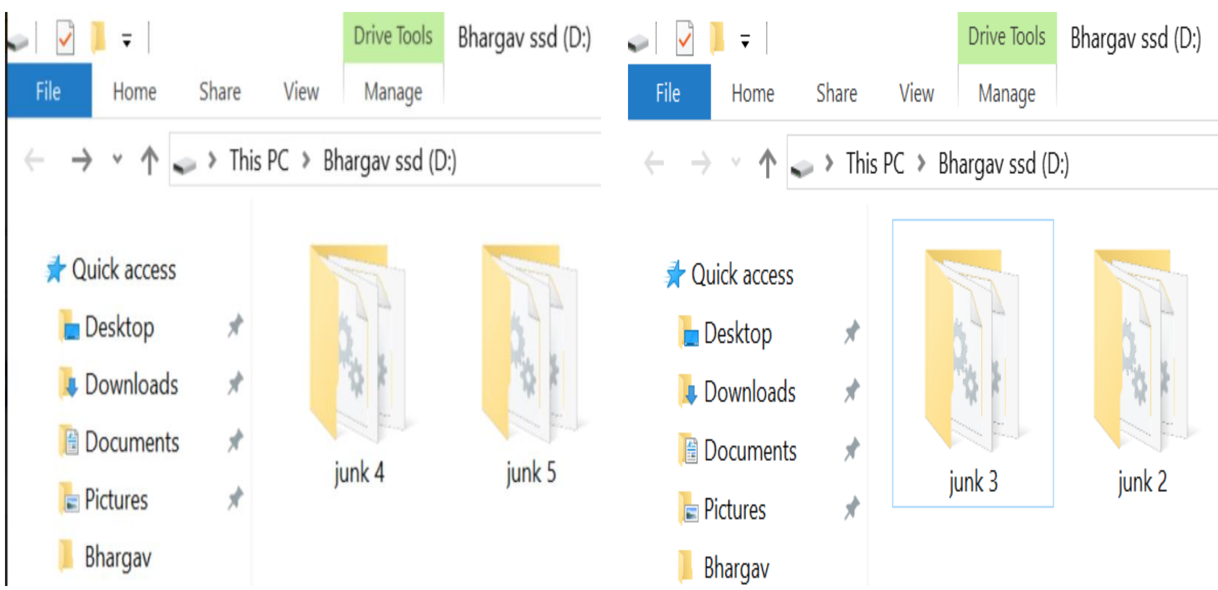

Figure 17. Passing Junk Folders in HDD and SSD

Installation of FTK Imager.

1. First, we go to https://accessdata.com/product-download/ftk-imager-version-4.2.0 and click on download now as follows:

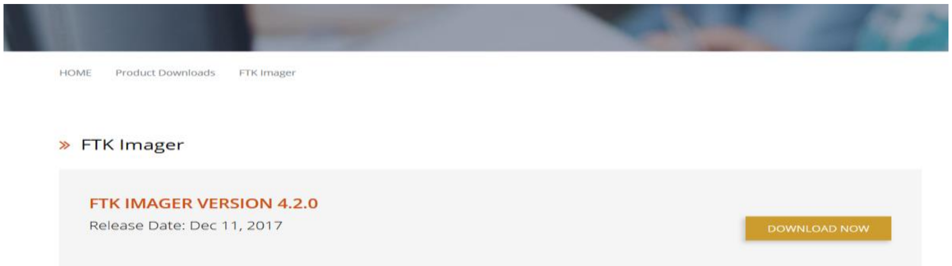

Figure 18. Webpage for Downloading FTK Imager

2. Then we need to fill in the form for getting the download link as follows:

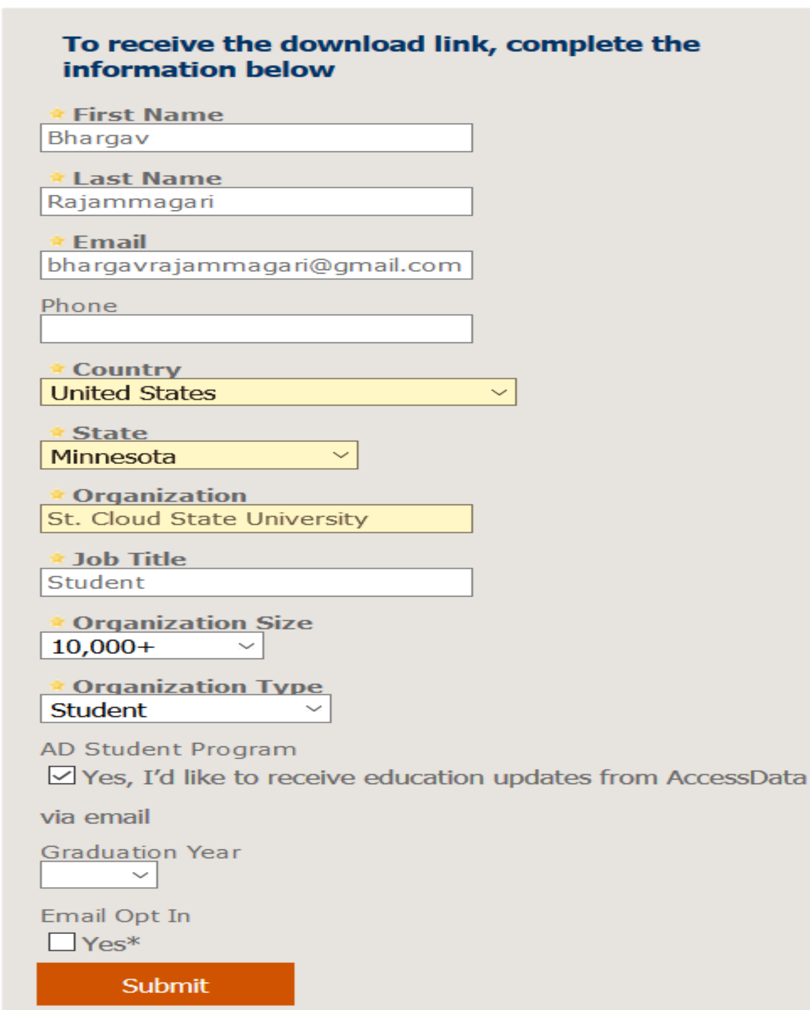

Figure 19. Registration Form for Downloading FTK Imager

3. We can see the link which we received in the email as follows:

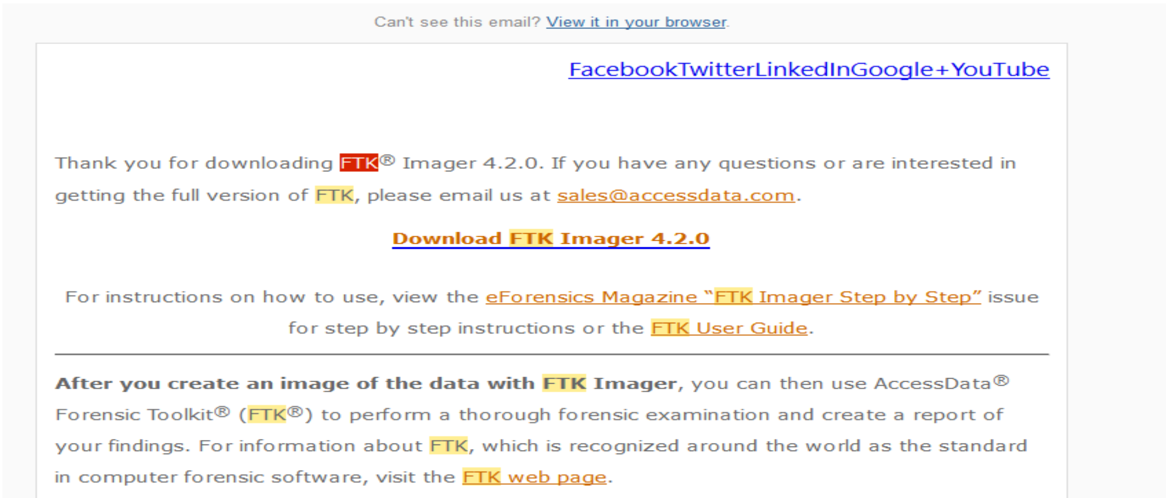

Figure 20. Email for Downloading FTK Imager

4. Installation Steps after downloading the FTK.exe file:

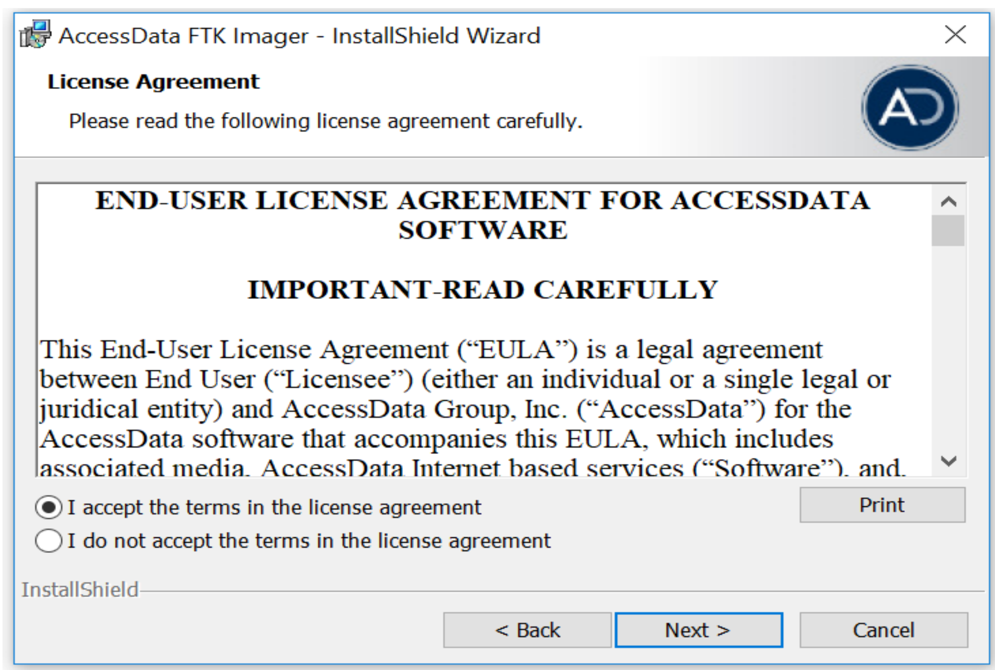

Figure 21. License Agreement for FTK Imager Download

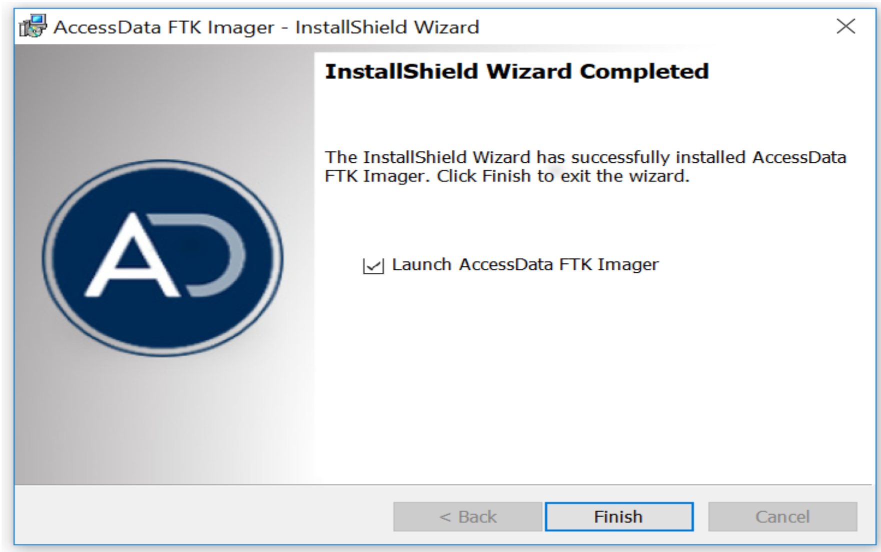

Figure 22. FTK Imager Installation Completed

Creating image of SSD. Now we open FTK imager and choose the source file for selecting the SSD. We can see the preview as follows:

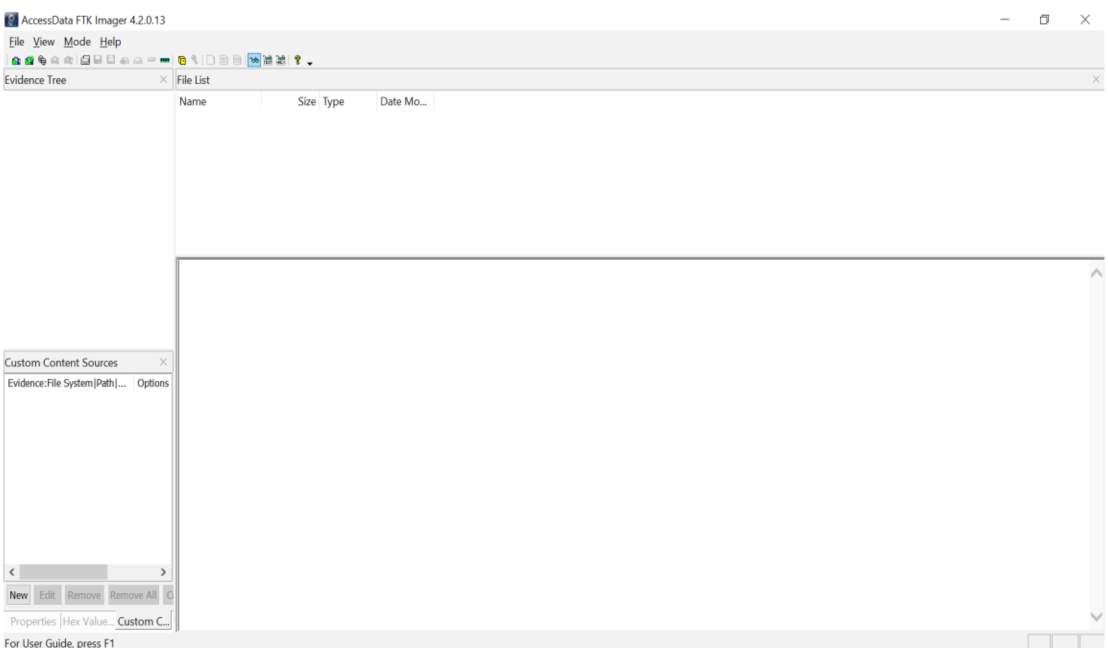

Figure 23. FTK Imager Welcome Page

Select the file and create a disk image option.

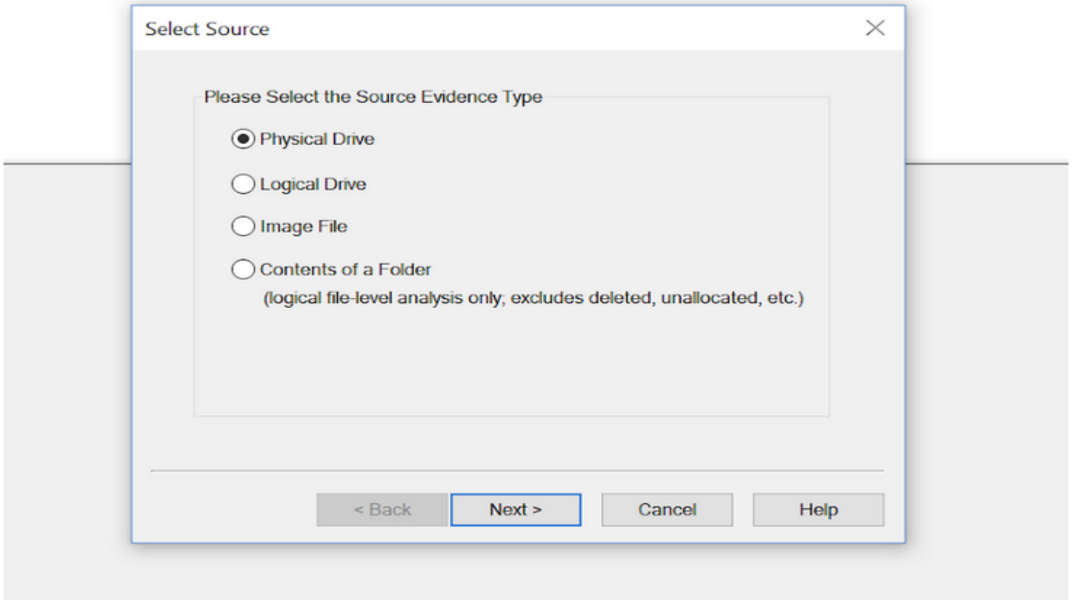

Figure 24. Selecting Source for FTK Imager

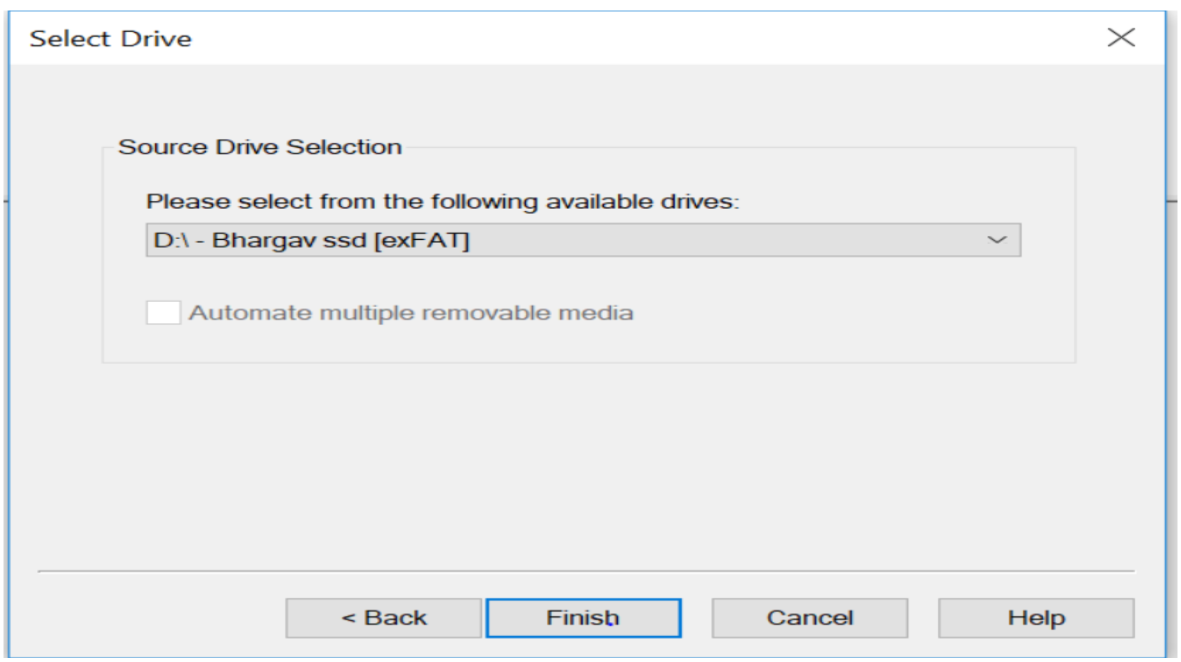

Figure 25. Selecting Drive for FTK Imager

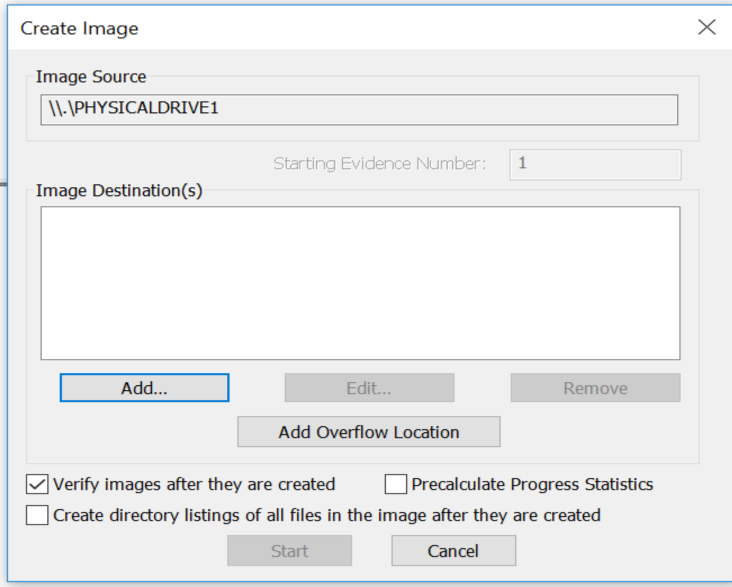

Figure 26. Selecting Image Source

Then we give in the evidence information for the image creation as follows:

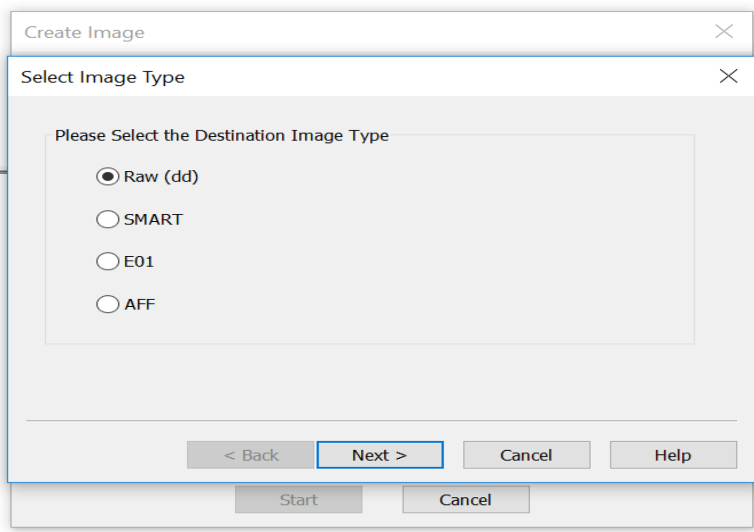

Figure 27. Selecting Image Type

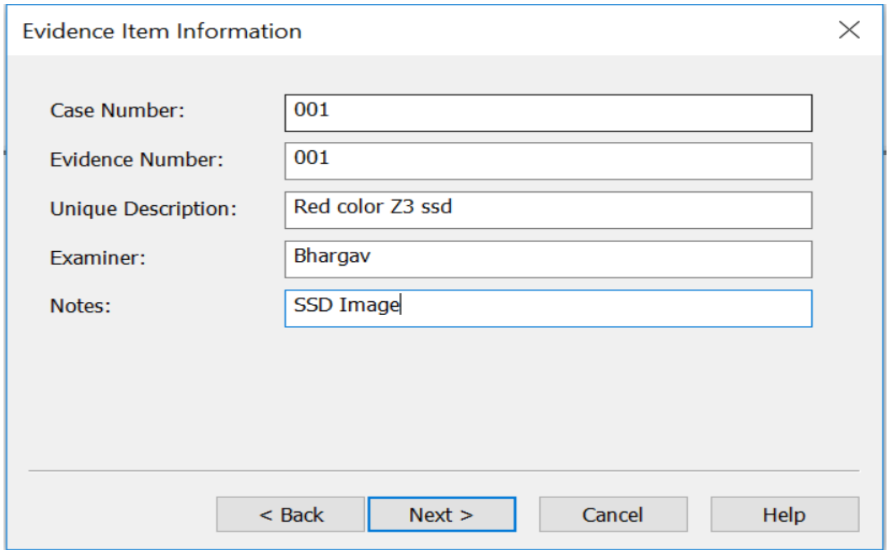

Figure 28. Evidence Information Form

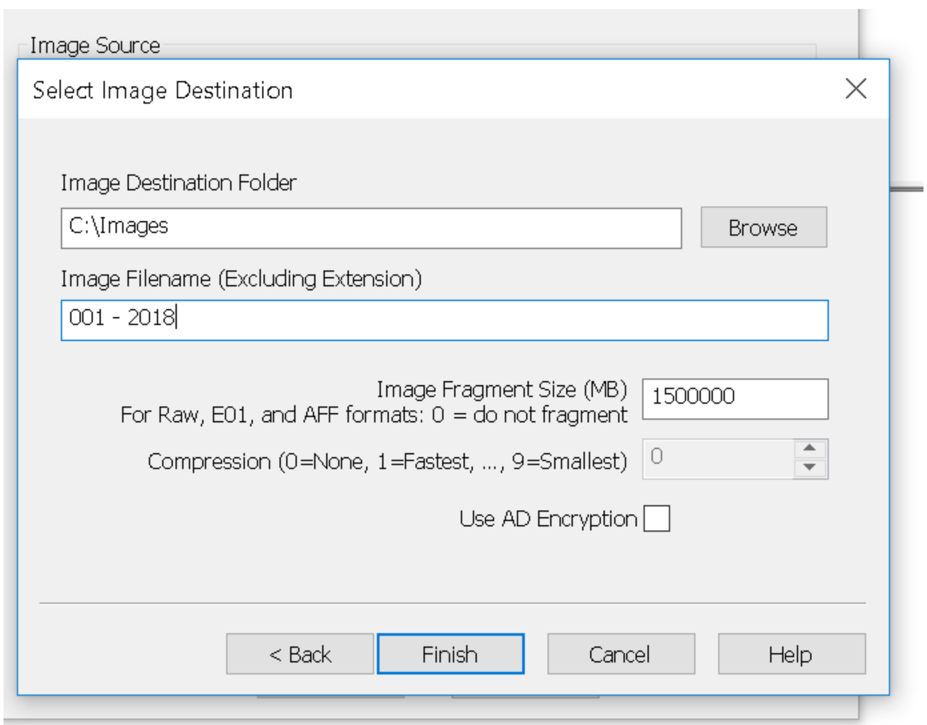

Figure 29. Selecting Image Destination Location

Then we can see the creation of an image as follows:

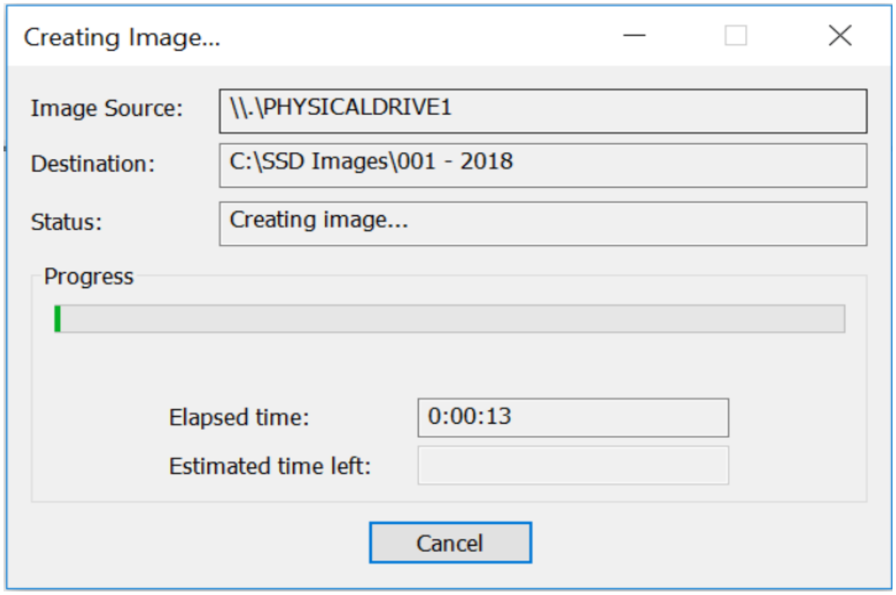

Figure 30. Creating Image

Then we can see that the image was created as follows:

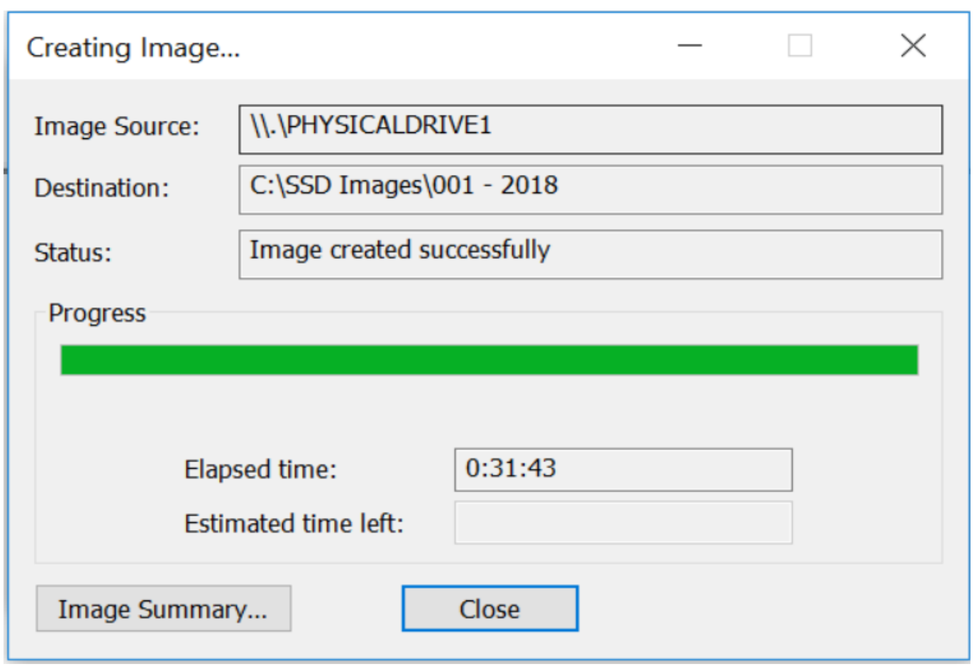

Figure 31. Image Created

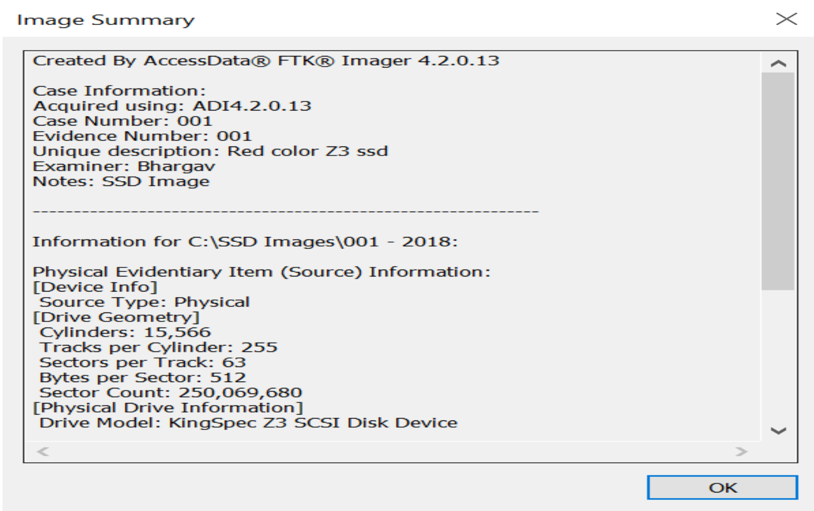

Figure 32. Summary of Image Created

SSD Image 2, Image 3, and Image 4 are constructed similarly.

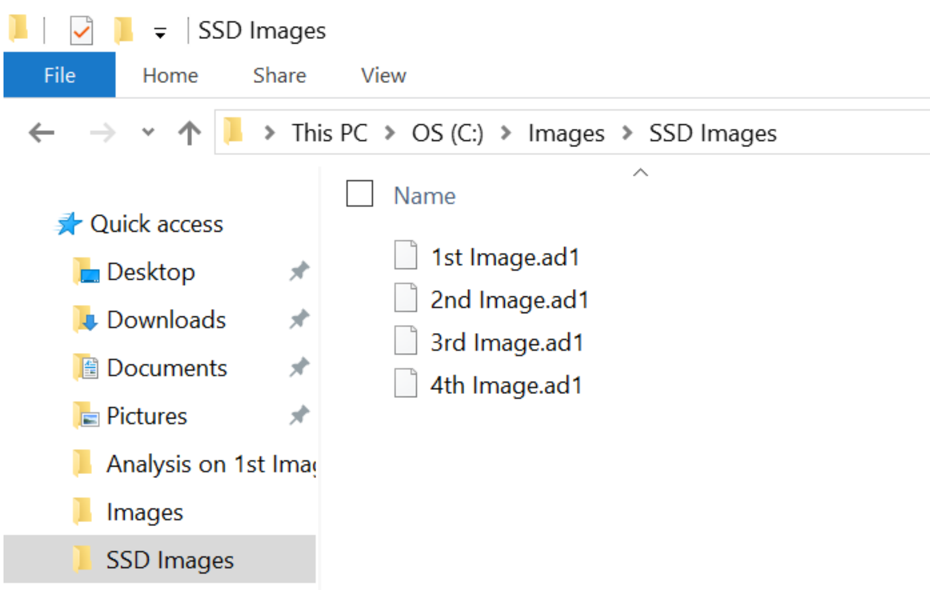

Figure 33. SSD Images in Folder

Creating image of HDD. As a part of the next process, we open the FTK Imager to create the image of the HDD. We select the “Source evidence type” as Physical Drive.

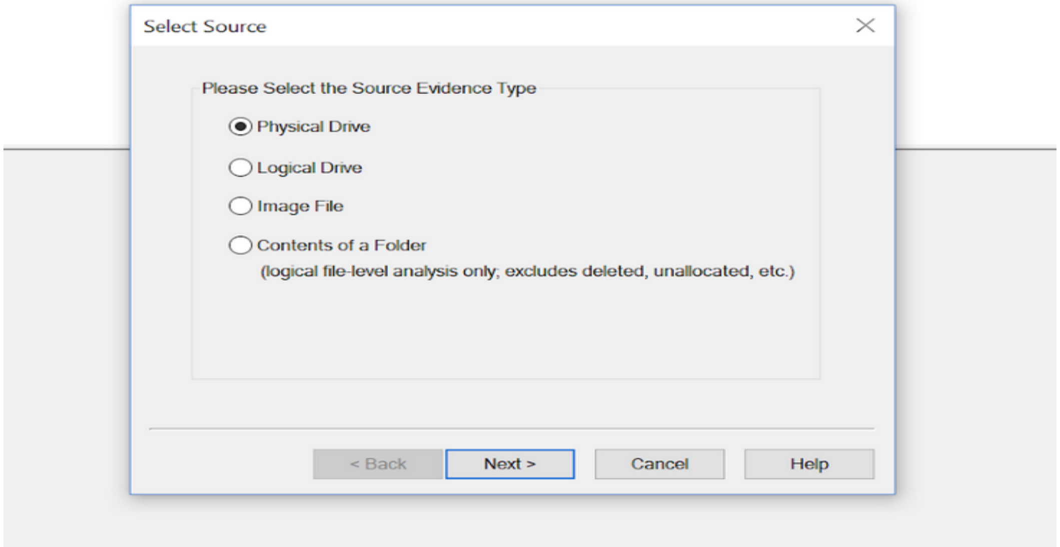

Figure 34. Selecting Source Type

Then in the next step, we choose the path for the available HDD, and we can see that in the image below:

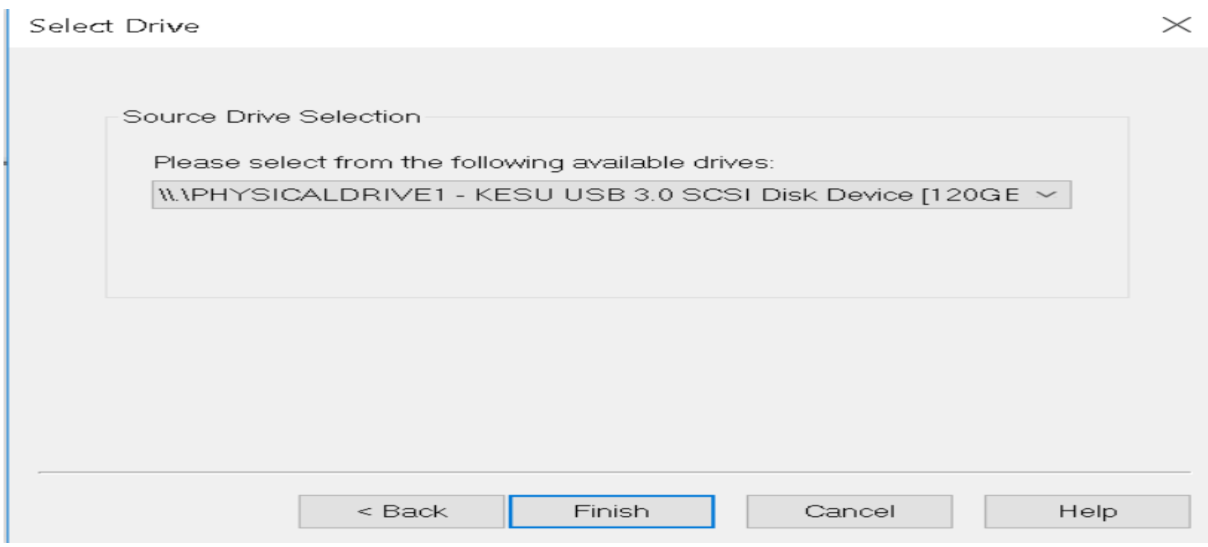

Figure 35. Selecting the Source Drive Selection

After selecting the drive, we click on finish, and this proceeds to the creation of an image of the HDD. We see the information what we enter for the examination purpose as follows:

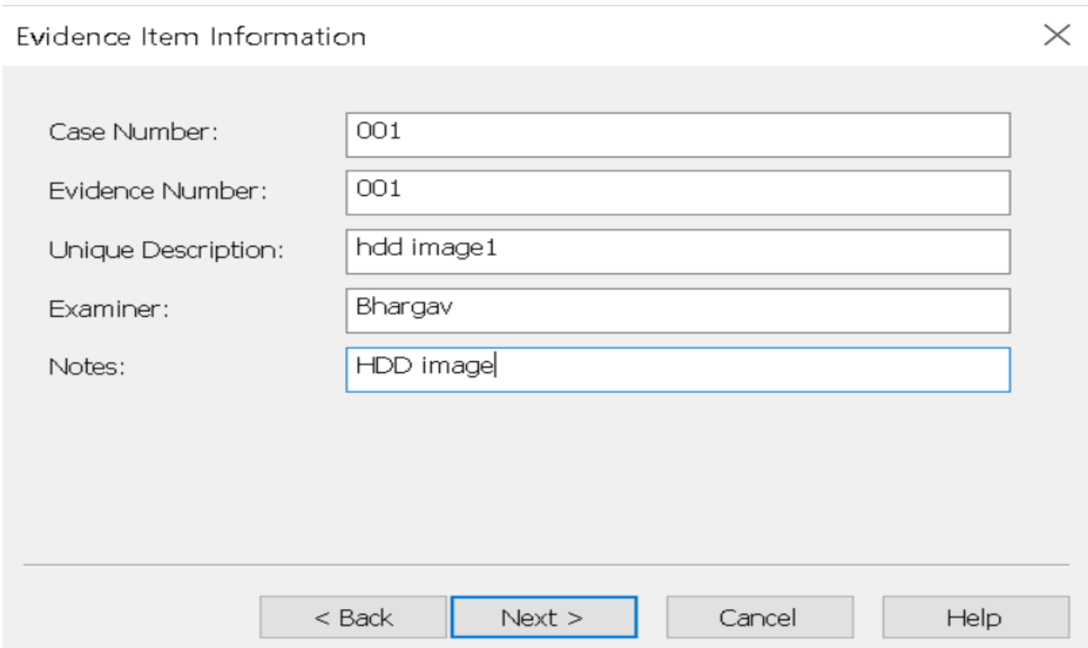

Figure 36. Evidence Item Information

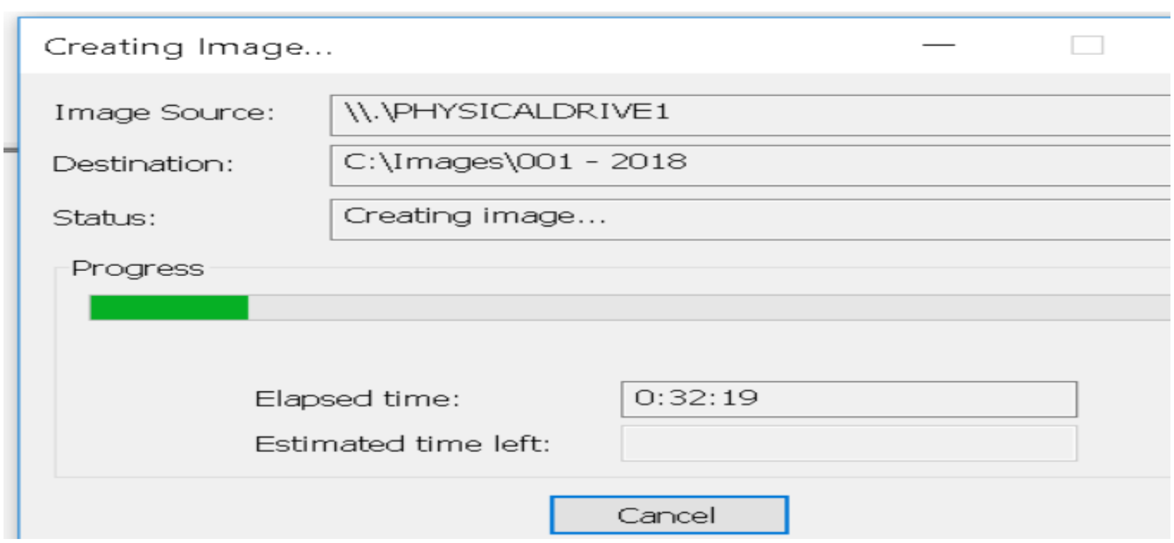

Figure 37. Creating SSD Image 1

As a part of the image creation, we can see the process as follows. Once it is done, we get an image of the HDD with all the contents of files and folders.

Installation of access data FTK suite. Access Data made forensic Tool Kit or FTK. It will scan a hard drive looking for different information. For example, to recover deleted emails, images. It also used to scan for text strings for using them as a password dictionary to crack the encryption.

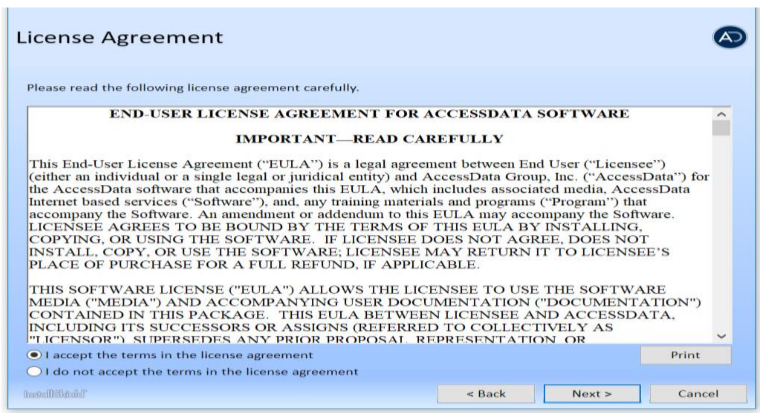

Figure 38. License Agreement for Access Data FTK Suite

After the installation:

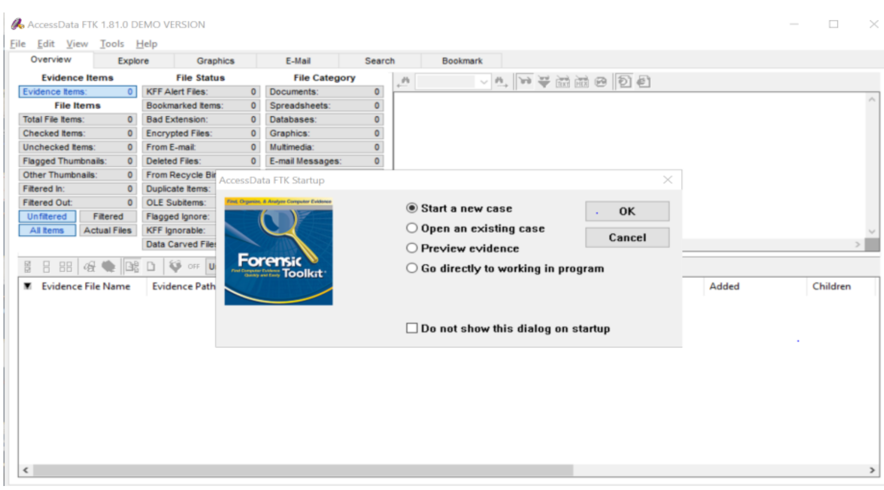

Figure 39. FTK Suite Welcome Page

To open and analyze a created image we need some software like FTK, Autopsy. In the below, I have a Forensic tool kit, and the First step is to select Start a New case and click OK. In the next level, we need to give the Investigator name. In this case, the investigator name is Bhargav. Case number refers to a unique number through which a case or investigation is assigned. The case name is the name of the case which we are about to investigate. Case path is giving the destination path, and case folder is the folder name to which the data must be stored.

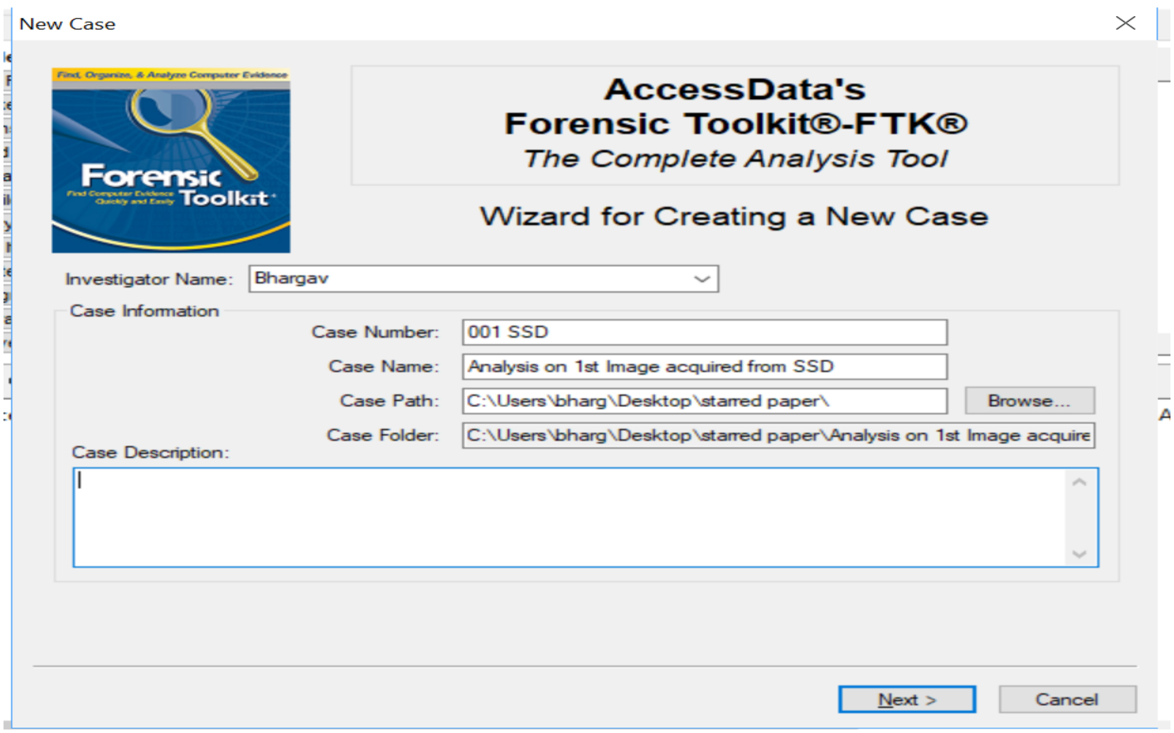

Figure 40. Creating a New Case

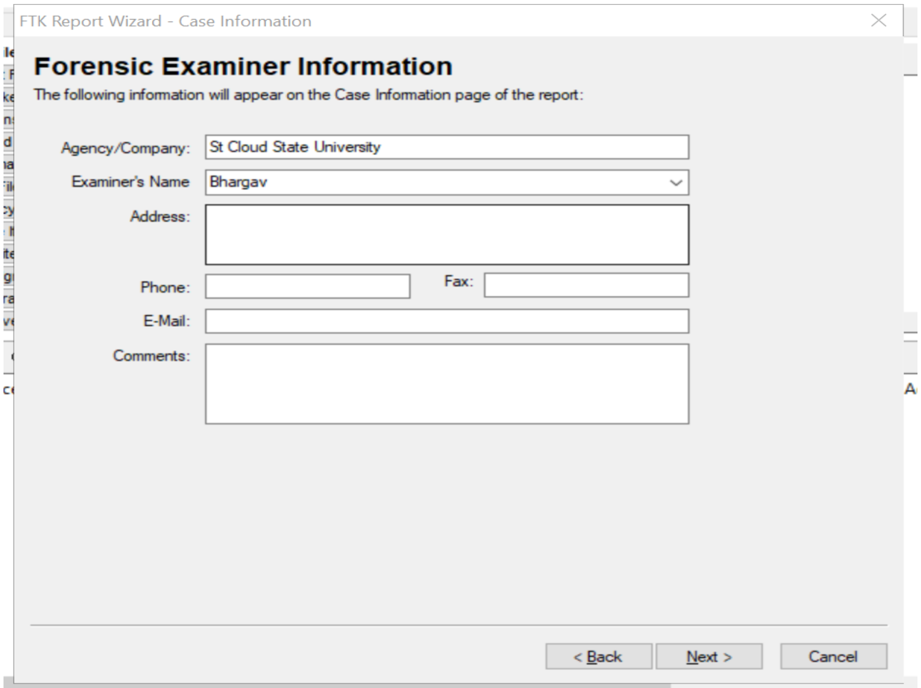

Figure 41. Forensic Examiner Information

To solve a case for a company or a university project, we need to mention that name. We must fill up the personal details like email or phone number address.

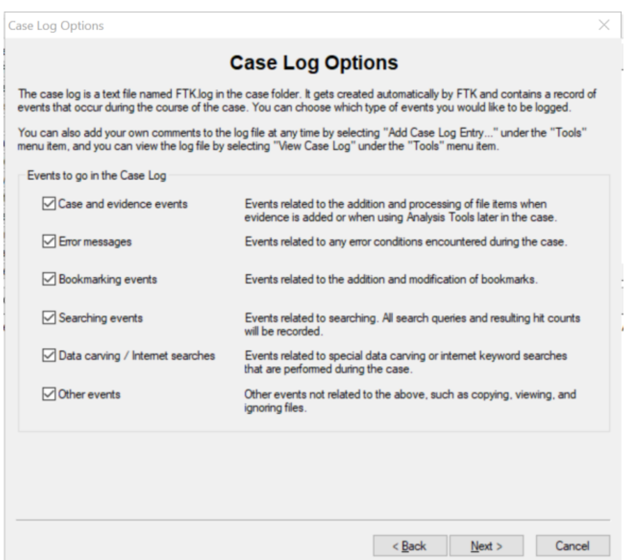

Figure 42. Case Log Options

These are the options that we must consider while performing a forensic investigation. By checking up the below fields, we can get the information about the errors and case events.

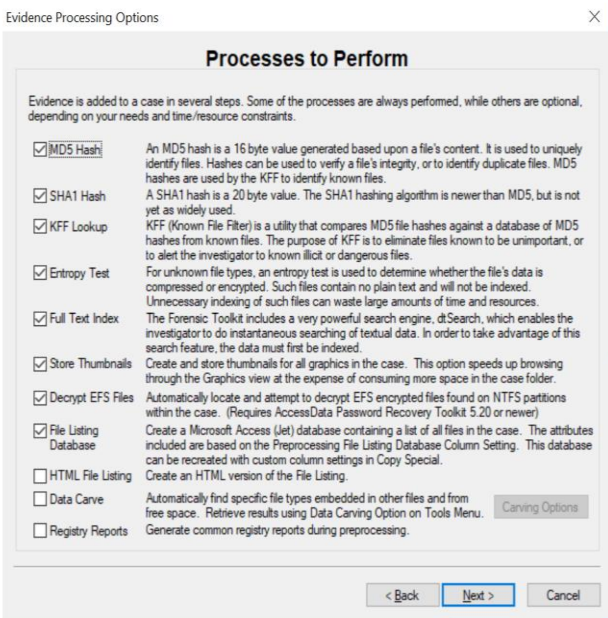

Figure 43. Processes to Perform

It helps in including the filters and what perspective are we supposed to get the results.

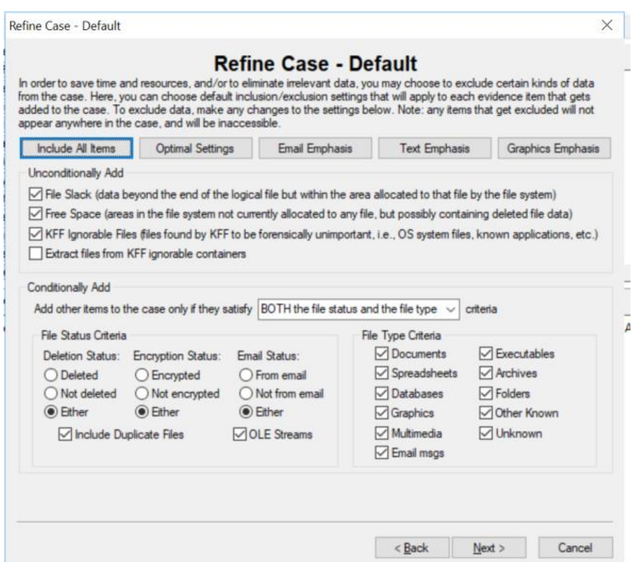

Figure 44. Default Case Setting

We need to check either we need deletion status or encryption status and email status.

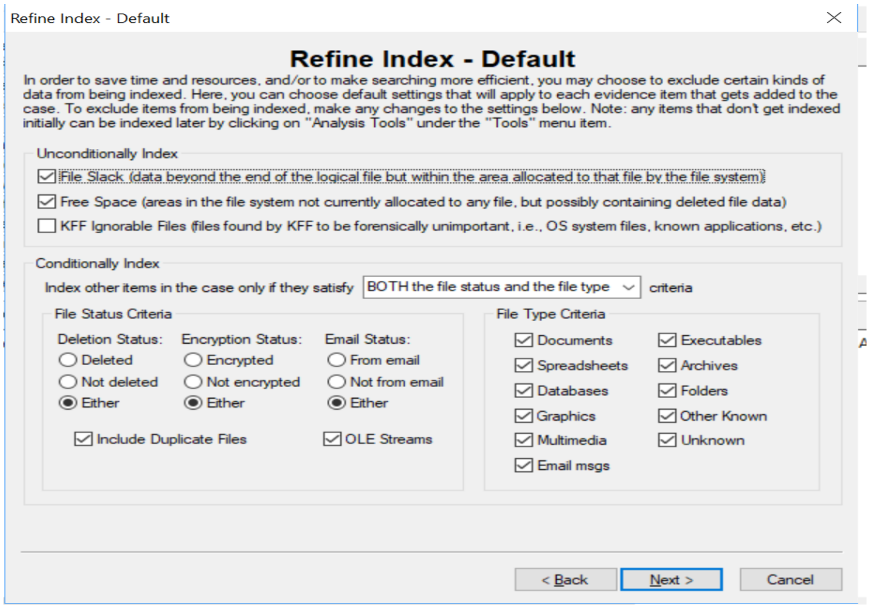

Figure 45. Default Index Setting

Checking the boxes like file slack helps in viewing the data inside the file slack as well as free spaces in the file system

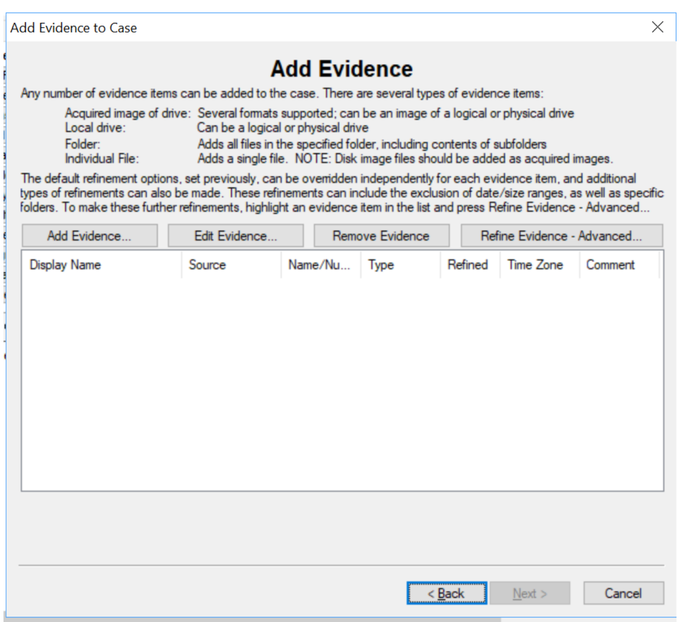

Figure 46. Adding Evidence

We can add the Created image as an evidence

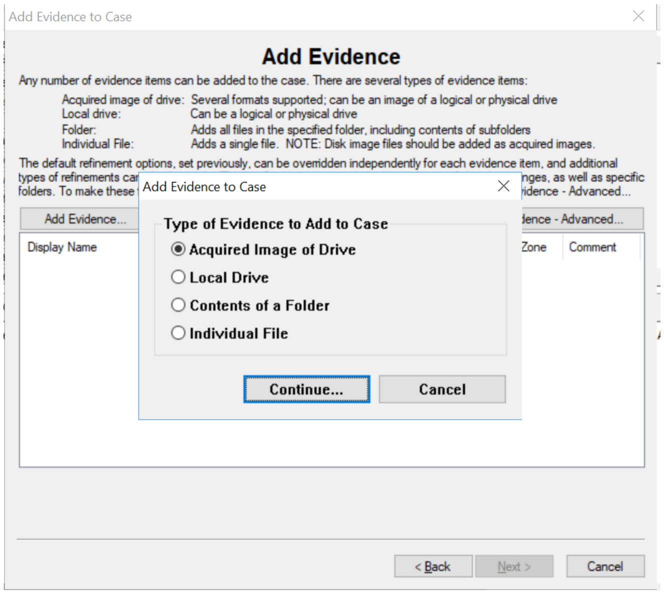

Figure 47. Type of Evidence to Add

Select the acquired image of the drive and select continue.

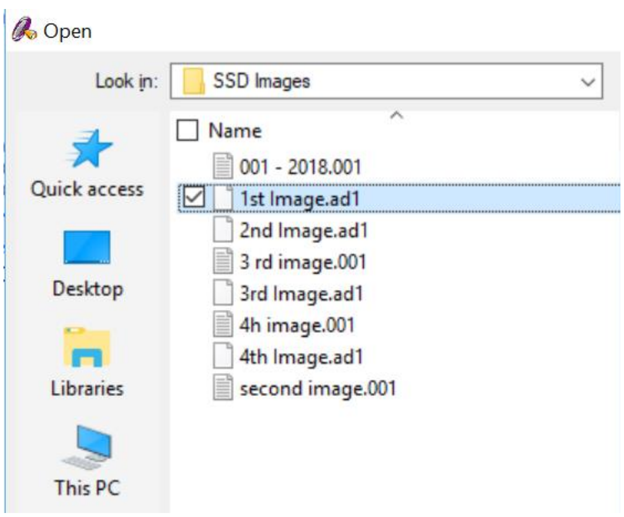

Figure 48. Selecting Acquired Image

Data Analysis

Select Image 1 and open the image.

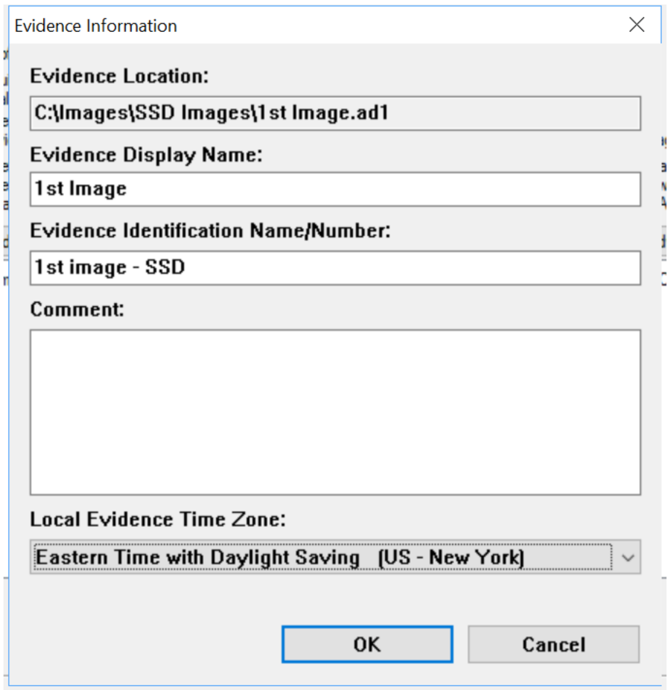

Figure 49. Evidence Information

Leave comments and the time zone we are staying

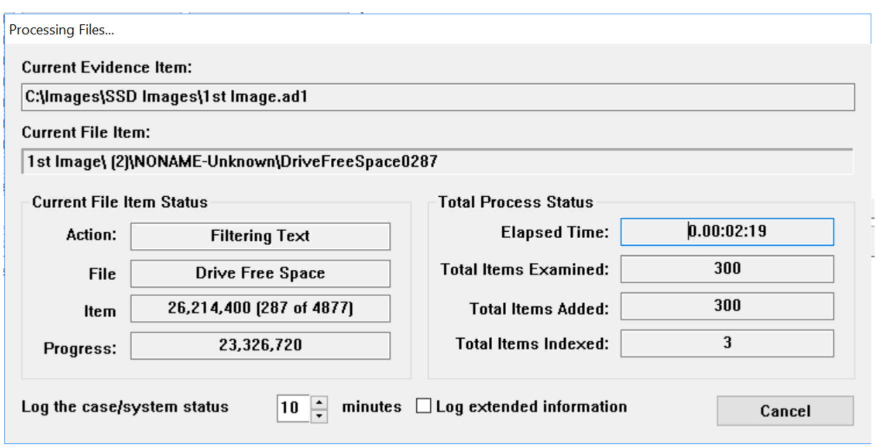

Figure 50. Leaving Time Zone when Processing an Image 1

This process will take approximately five to six hours to analyze the image.

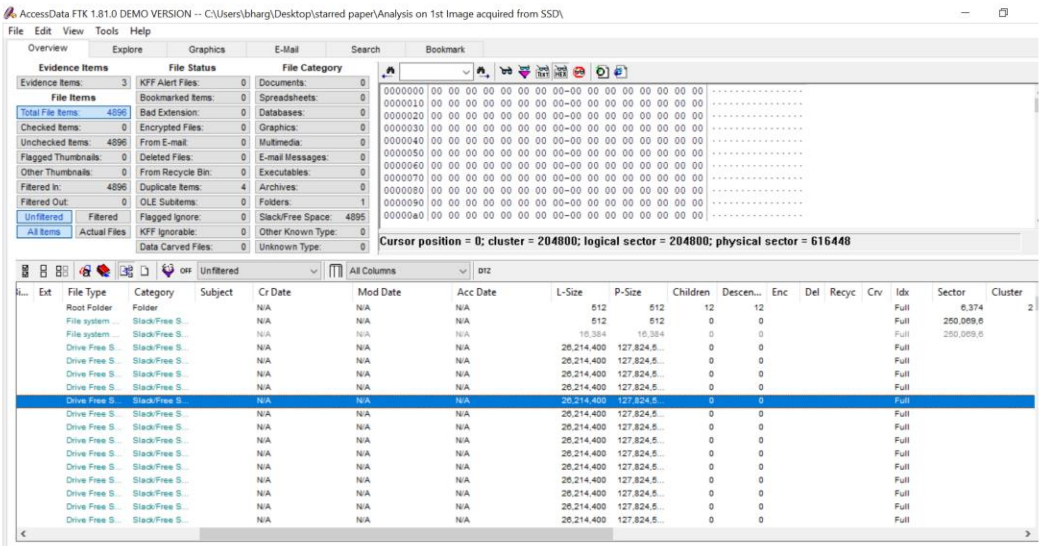

Figure 51. First Image Results

Results of the first image. Now in the Search Term box, we search for indexed Words. From the below image we can see a sample number of hits for the search items animals, travel, plant, and sunshine.

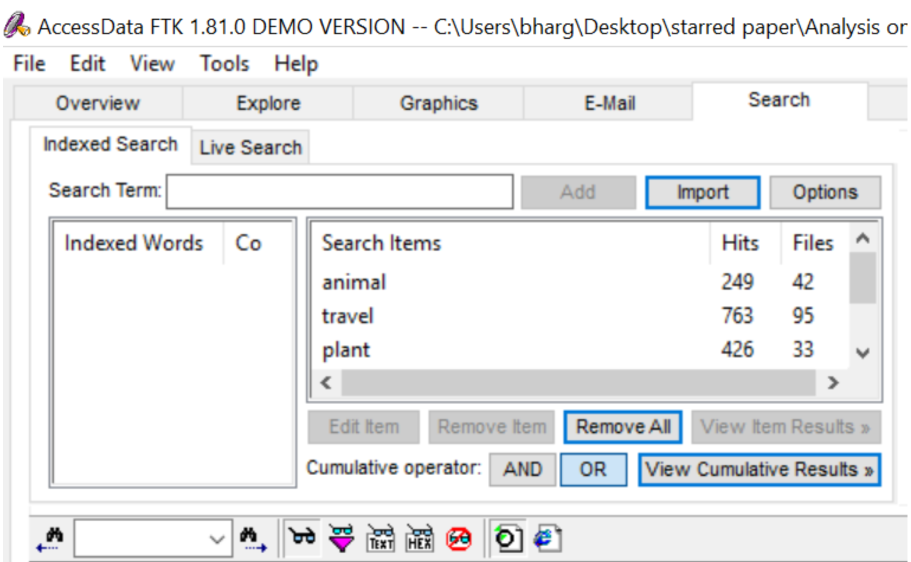

Figure 52. Sample Hits for Animals, Travel, and Plant for SSD Image 1

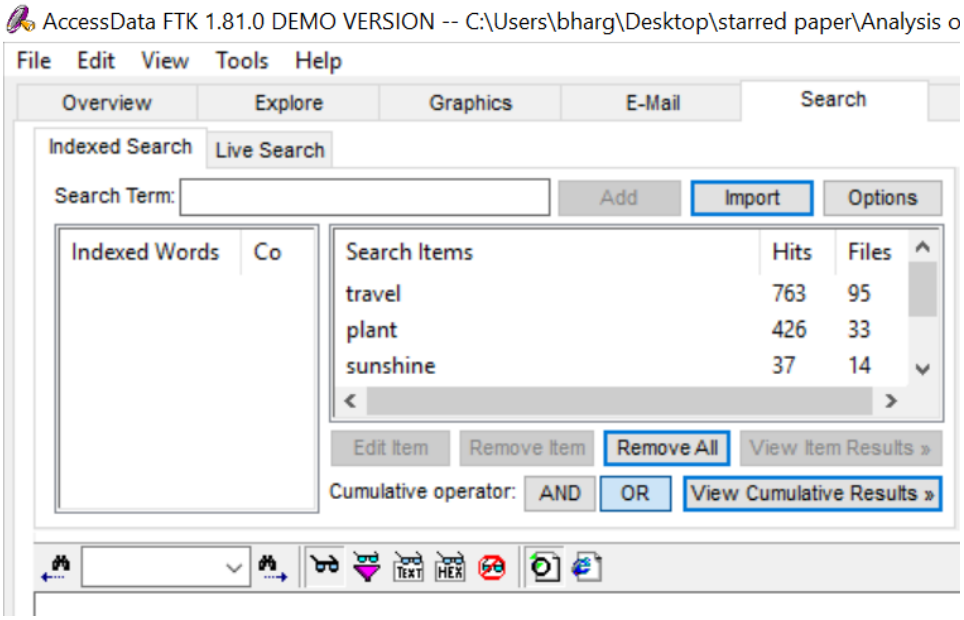

Figure 53. Sample Hits for Sunshine, Travel, and Plant for SSD Image 1

Now we search for .doc, .pdf, and .xlxs files, and we see the number of hits as follows.

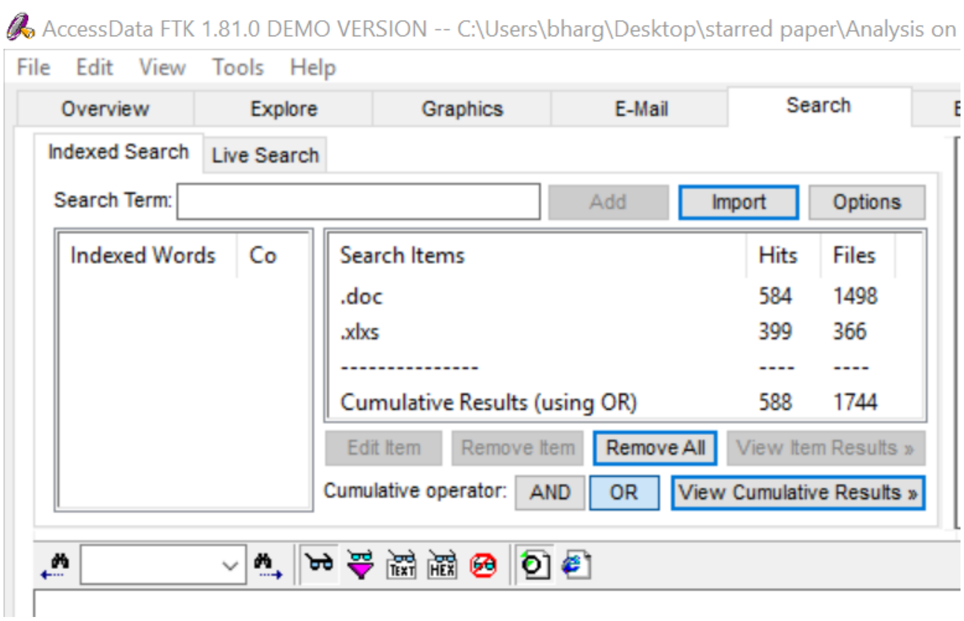

Figure 54. Sample Hits for .Doc and .Xlxs for SSD Image 1

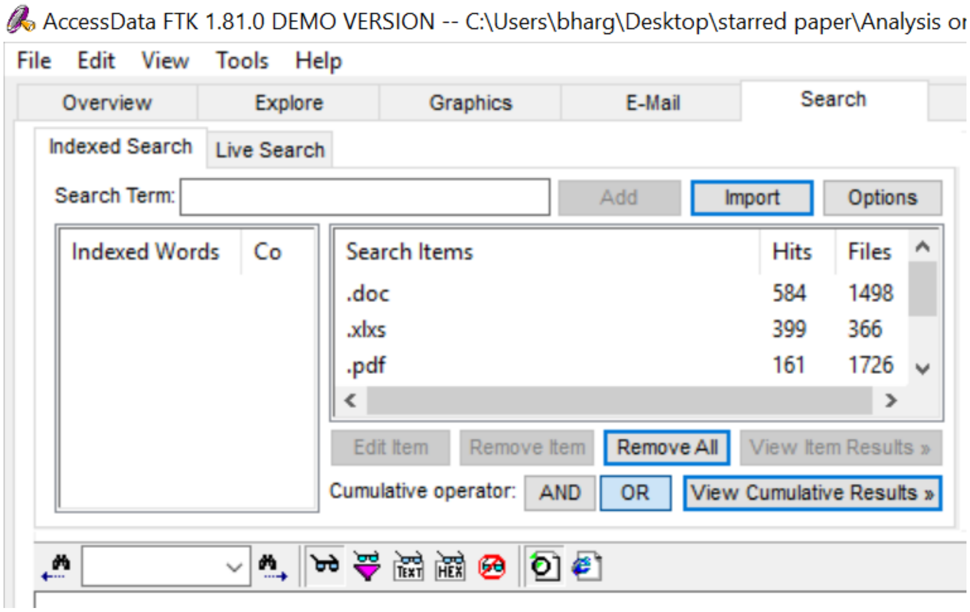

Figure 55. Sample Hits for .Doc, .Pdf, and .Xlxs for SSD Image 1

Analysis for second image of SSD. Now we fill in the case information such as Case Number, Case Name, Case Path, and Case Folder and click on next.

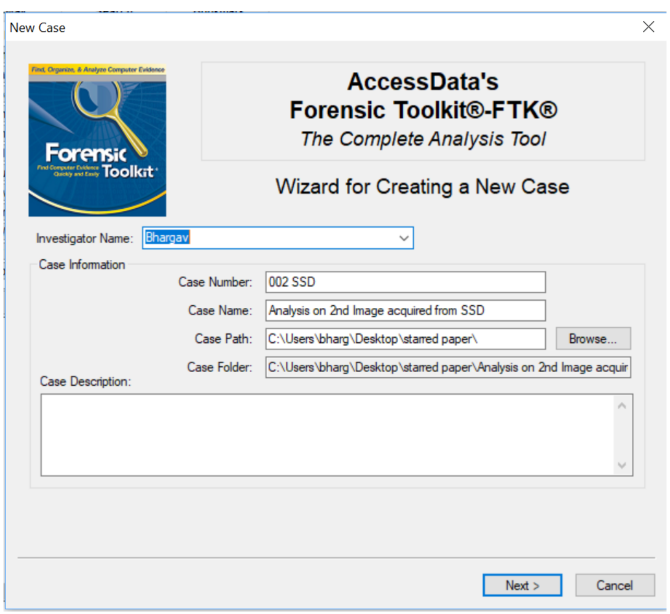

Figure 56. Creating a New Case for SSD Image 2

Now when we click on the overview, we can see 4896 items.

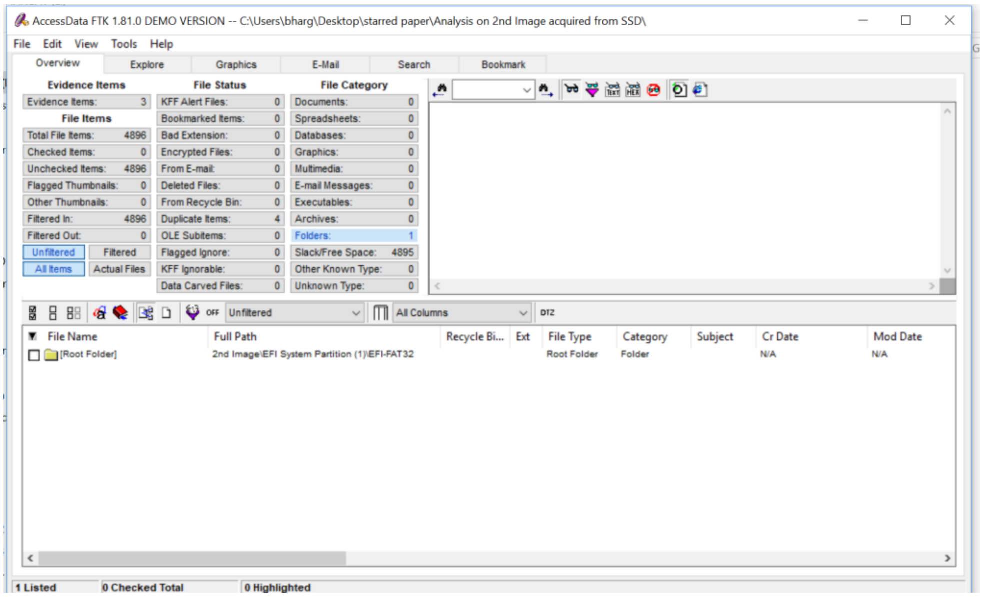

Figure 57. Overview of SSD Second Image

We can see the Hex code format for one of the selected files as follows:

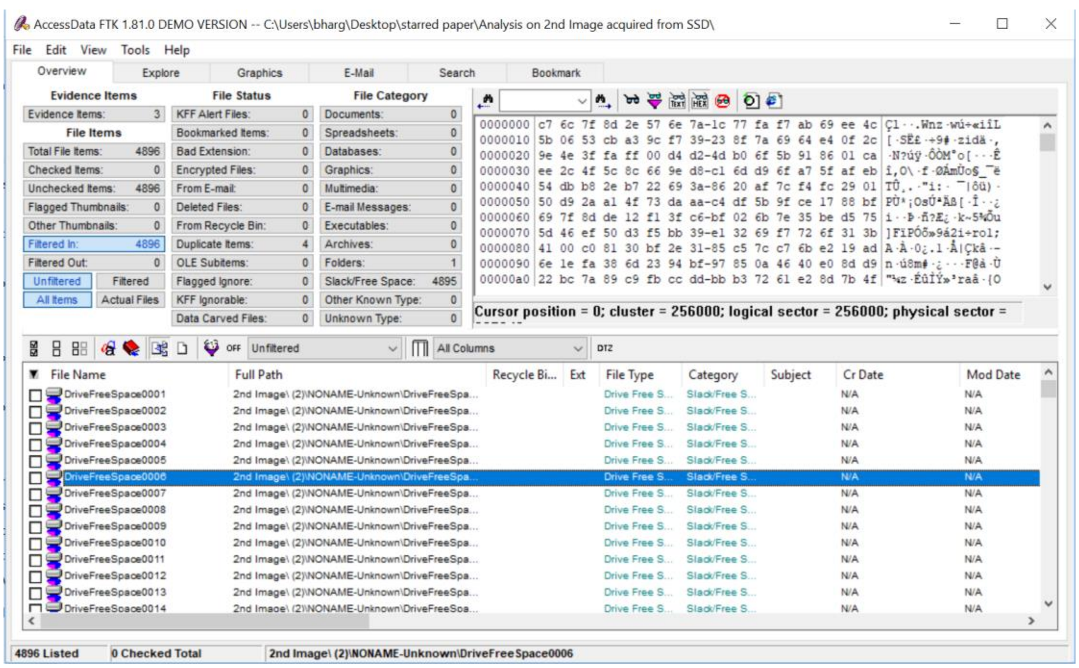

Figure 58. Hex Code of SSD Second Image

Second image results. Now in the Search Term box, we search for indexed Words. From the below image we can see a sample number of hits for the search items animals, travel, plant, and sunshine.

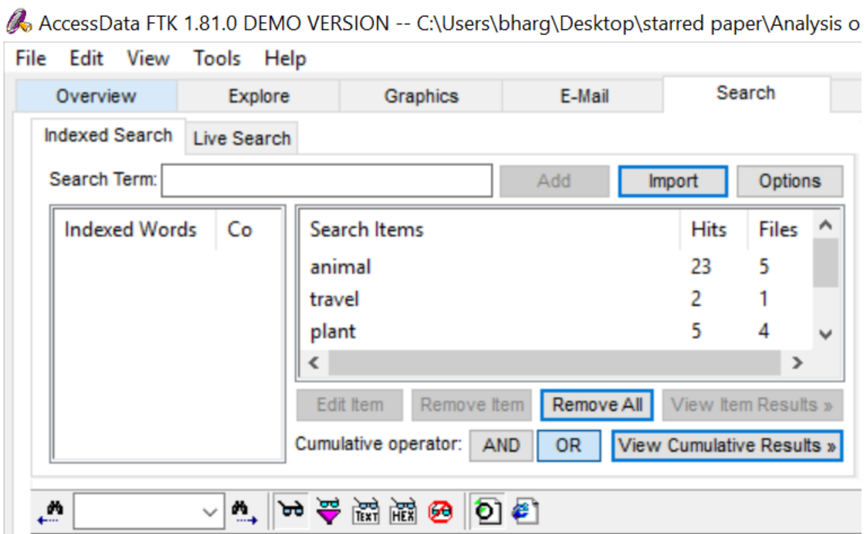

Figure 59. Sample Hits for Animals, Travel, and Plant for SSD Image 2

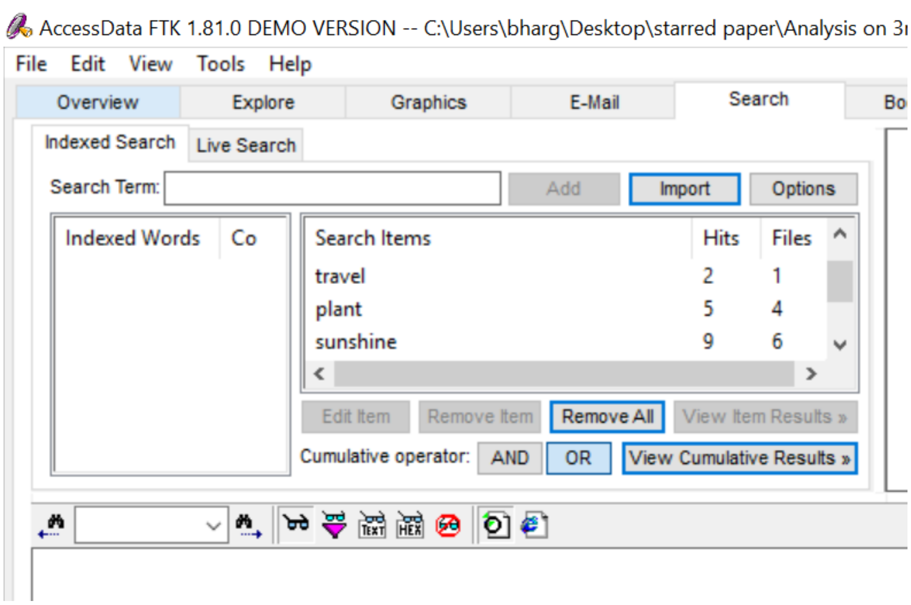

Figure 60. Sample Hits for Sunshine, Travel, and Plant for SSD Image 2

Now we search for .doc, .pdf and .xlxs files, and we see the number of hits as follows.

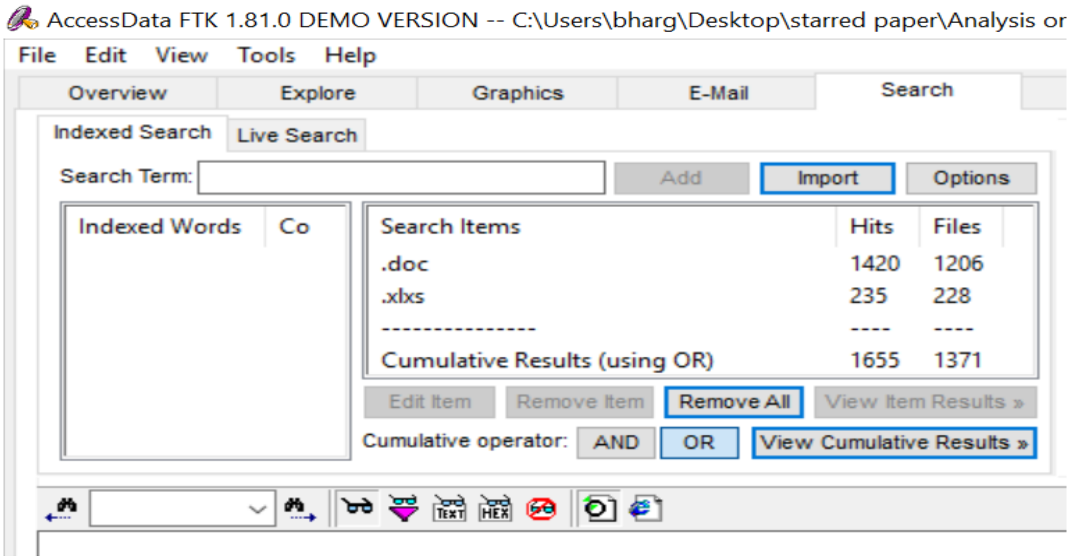

Figure 61. Sample Hits for .Doc and .Xlxs for SSD Image 2

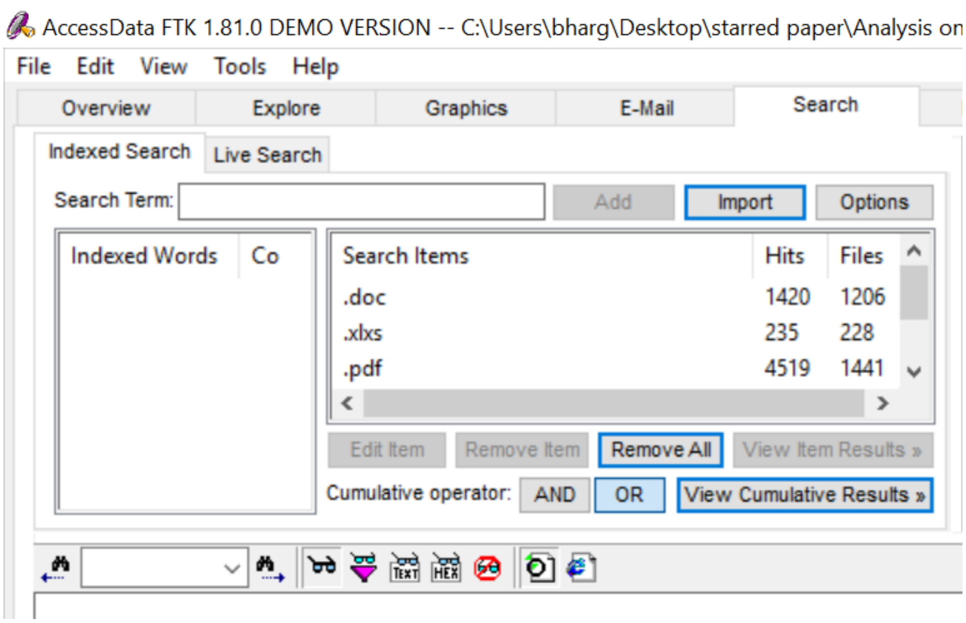

Figure 62. Sample Hits for .Doc, .Pdf, and .Xlxs for SSD Image 2

Now we fill in the case information such as Case Number, Case Name, Case Path, and Case Folder and click on next.

Analysis of SSD third image.

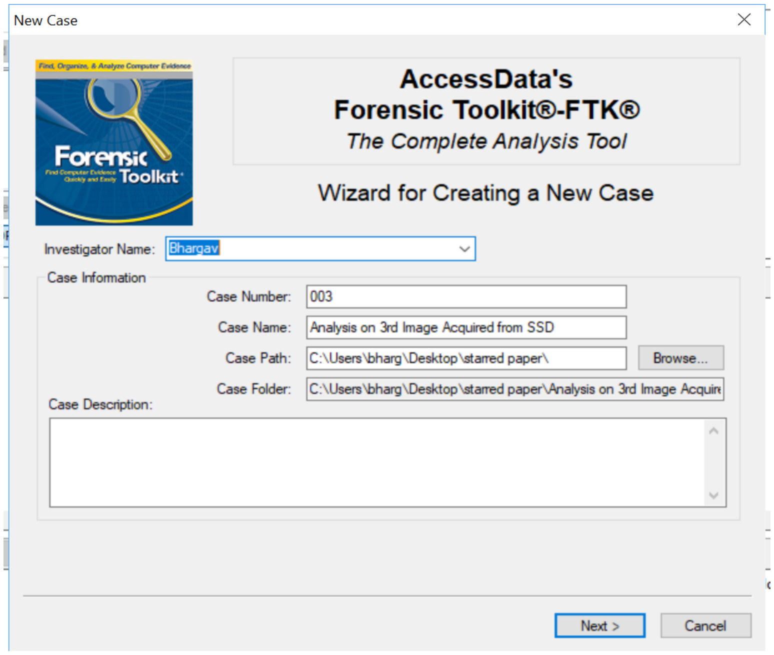

Figure 63. Creating a New Case for SSD Image 3

Now we fill the evidence information, such as Evidence location, Evidence Display Name and click on OK.

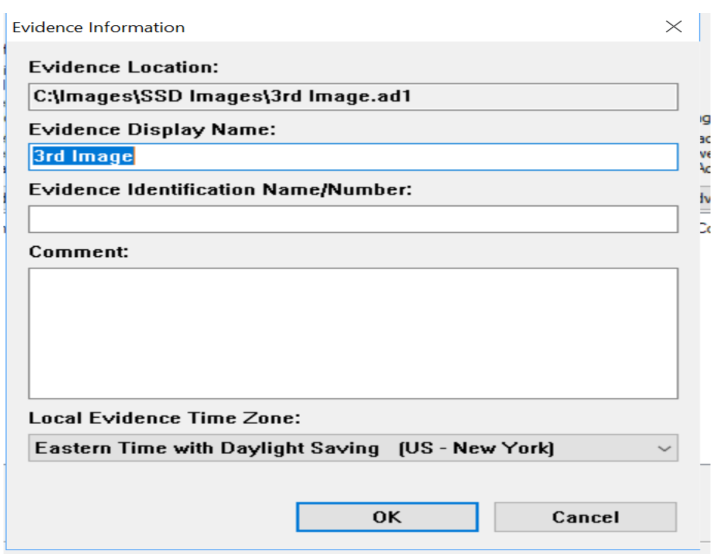

Figure 64. Evidence Information for SSD Third Image

Now we see the status of the image analysis.

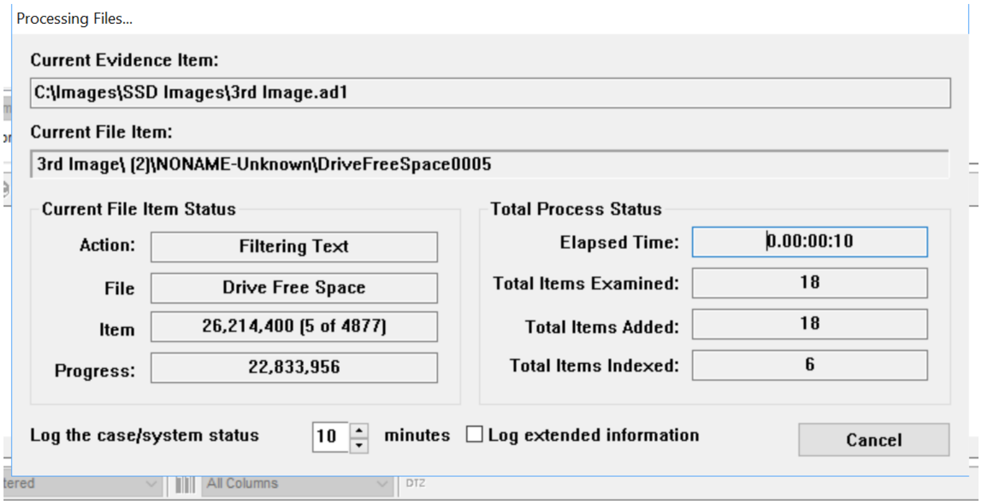

Figure 65. Leaving Time Zone when Processing an Image 2

Results of SSD third image. After the image analysis is done, we can see the number of files. When we click on one of the files, we can see the Hex code format as follows:

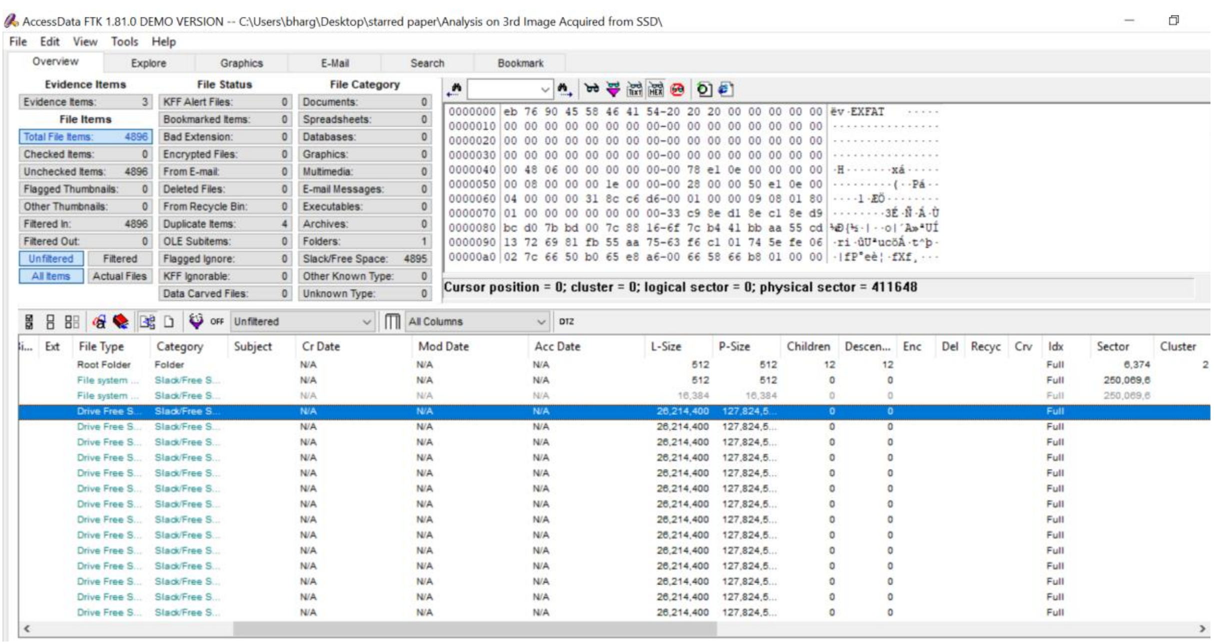

Figure 66. Hex Code for SSD Third Image

Now in the Search Term box, we search for indexed Words. From the below image we can see a sample number of hits for the search items animals, travel, plant, and sunshine.

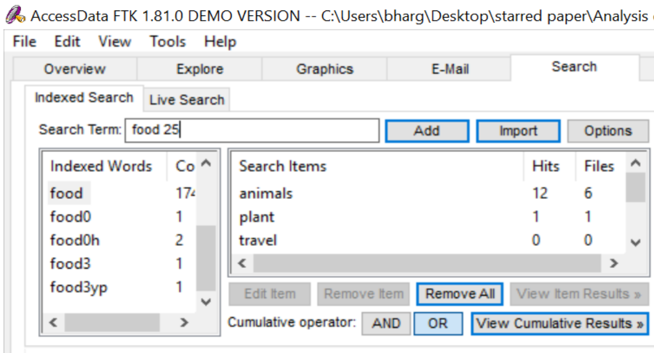

Figure 67. Sample Hits for Animals, Travel, and Plant for SSD Image 3

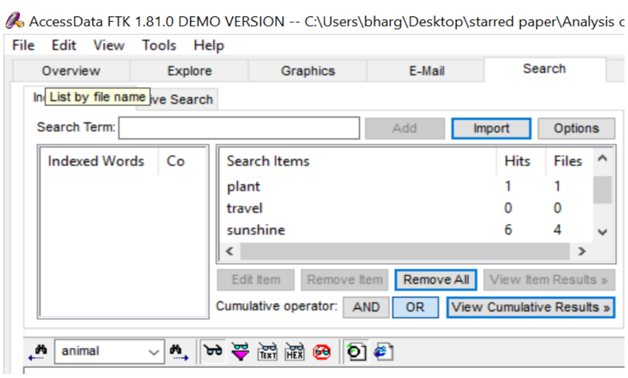

Figure 68. Sample Hits for Sunshine, Travel, and Plant for SSD Image 3

Now we search for .doc, .pdf and .xlxs files, and we see the number of hits as follows.

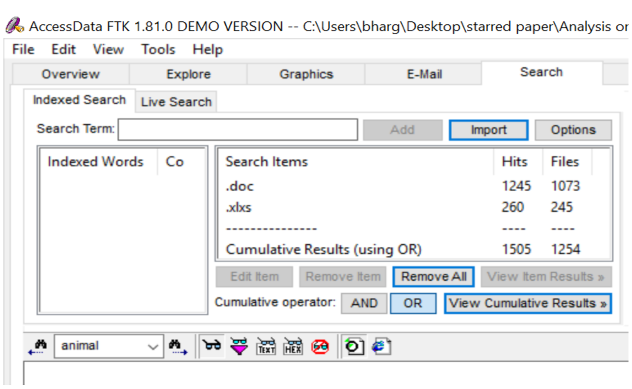

Figure 69. Sample Hits for .Doc and .Xlxs for SSD Image 3

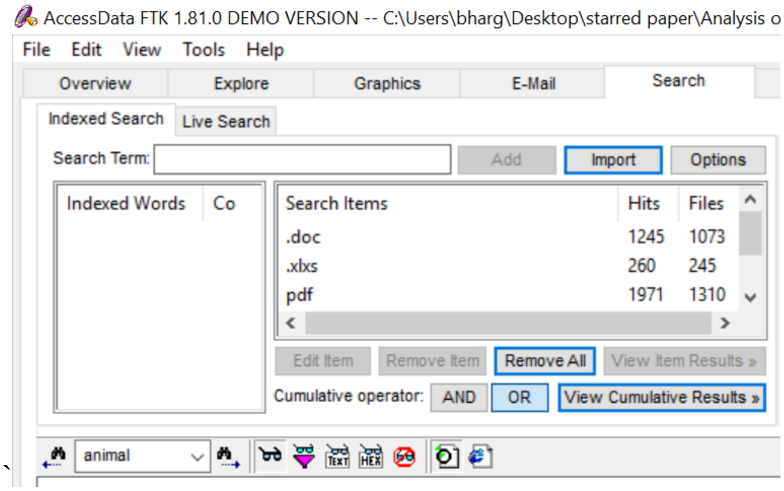

Figure 70. Sample Hits for .Doc, .Pdf, and .Xlxs for SSD Image 3

Analysis of SSD fourth image. Now we fill in the case information such as Case Number, Case Name, Case Path, and Case Folder and click on next.

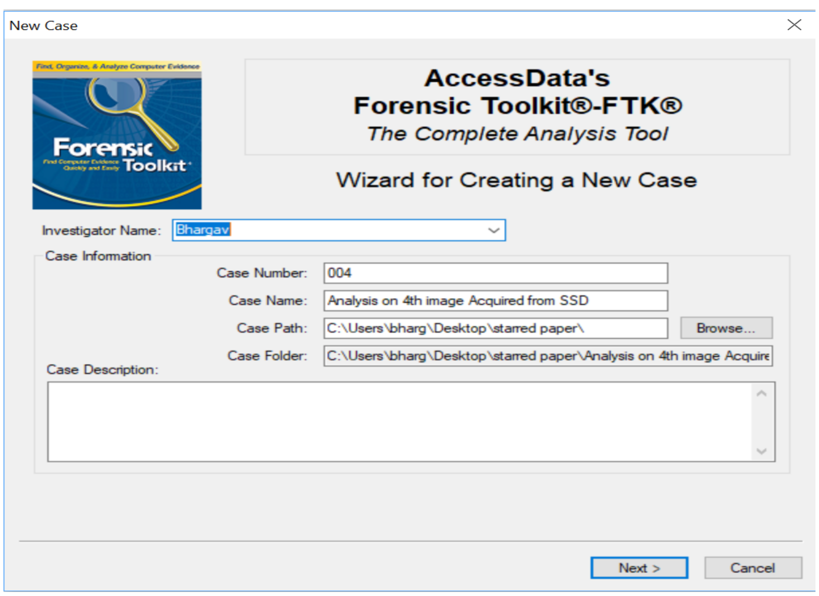

Figure 71. Creating a New Case for SSD Image 4

Now we fill the evidence information, such as Evidence location, Evidence Display Name and click on OK.

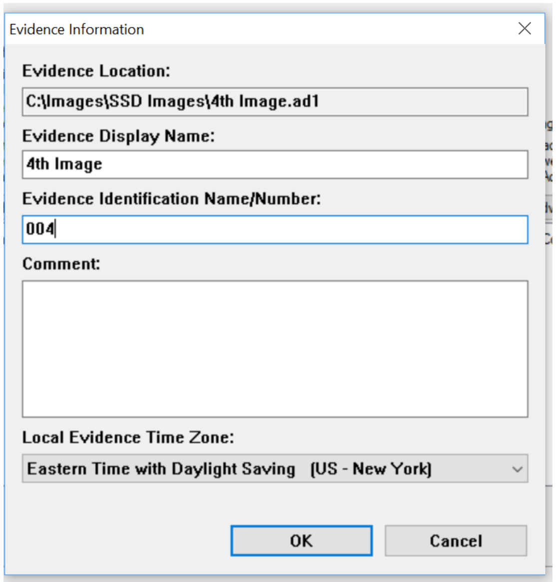

Figure 72. Evidence Information for SSD Image 4

Now we see the status of the image analysis.

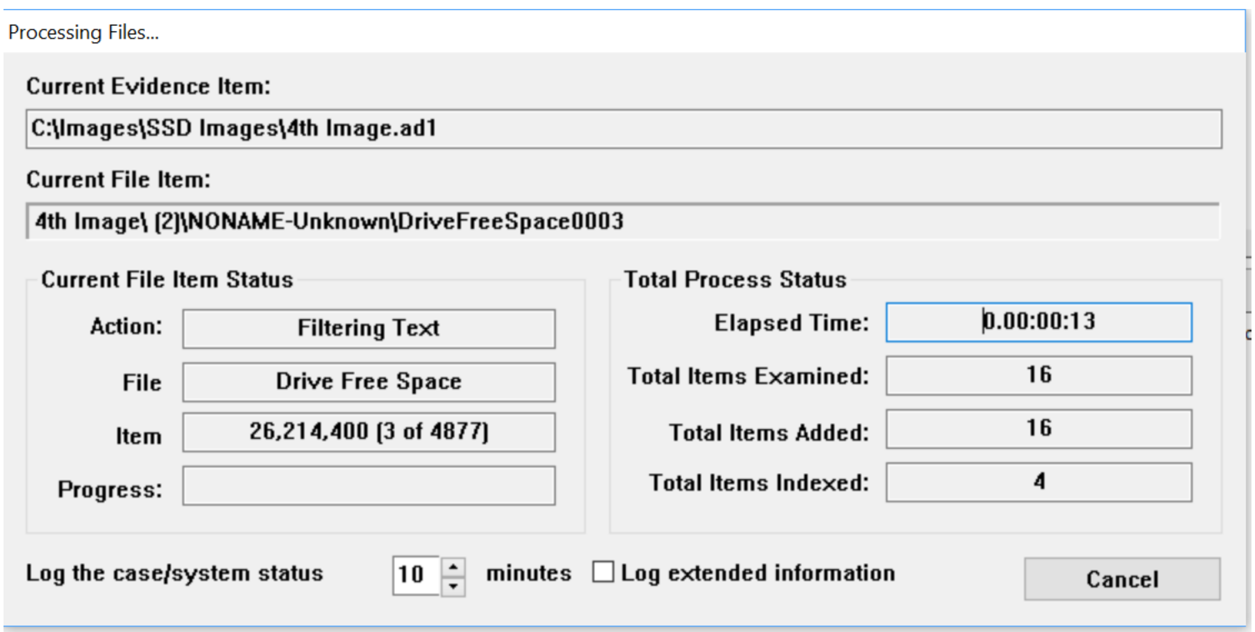

Figure 73. Leaving Time Zone when Processing an Image 4

After the image analysis is done, we can see the number of files. When we click on one of the files, we can see the Hex code format as follows:

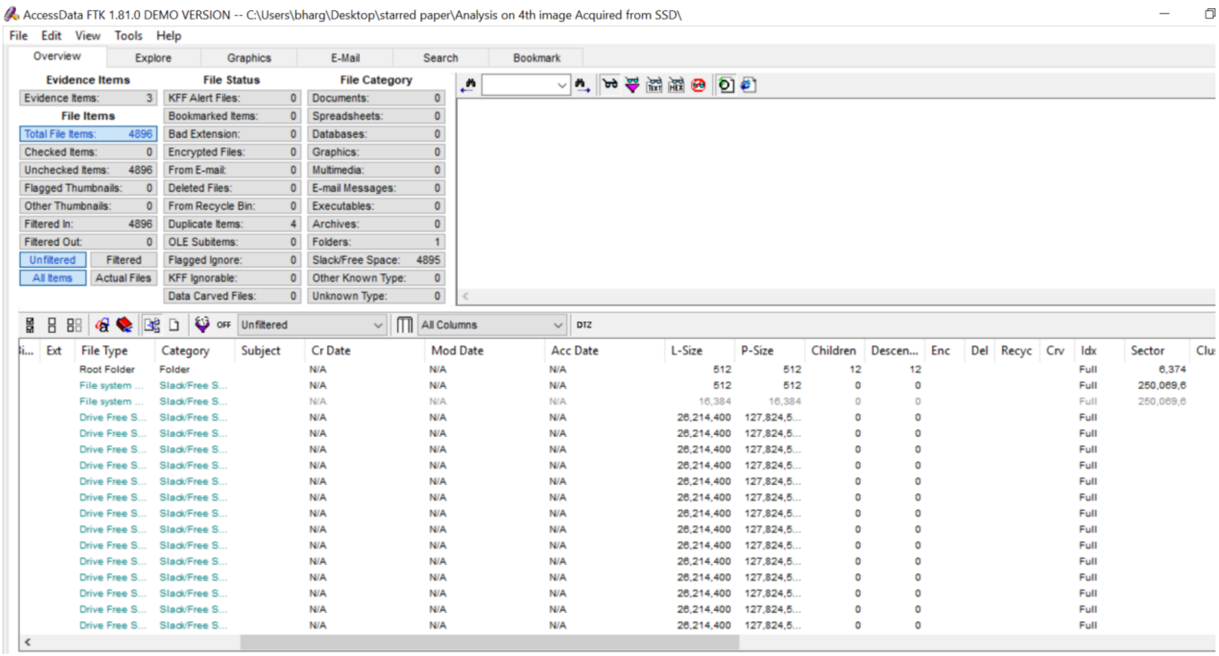

Figure 74. Hex Code for SSD Fourth Image

Now in the Search Term box, we search for indexed Words. From the below image we can see a sample number of hits as 0 for the search items animals, travel, plant, and sunshine.

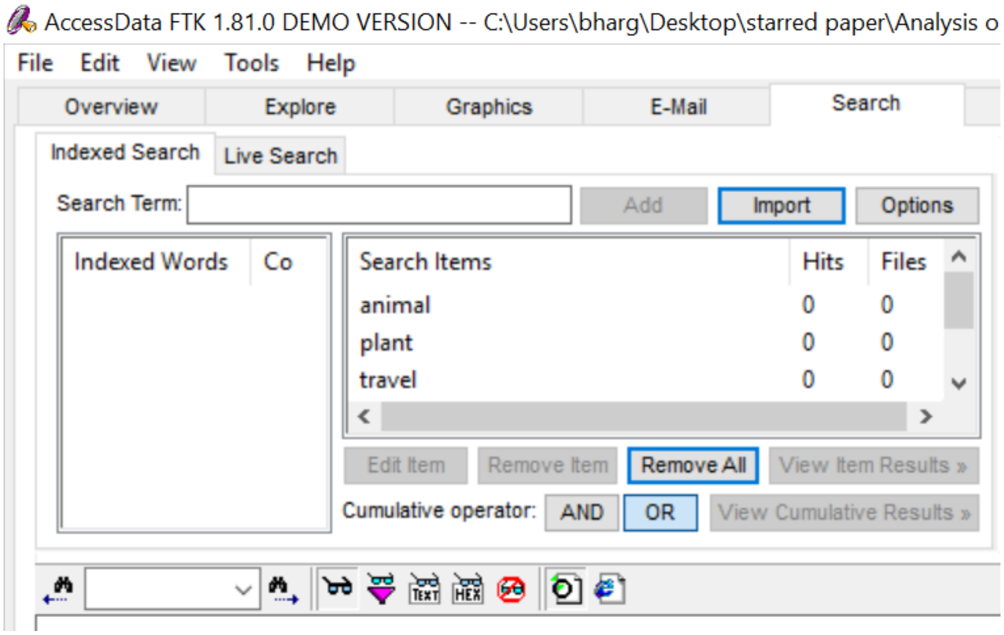

Figure 75. Sample Hits for Animals, Travel, and Plant for SSD Image 4

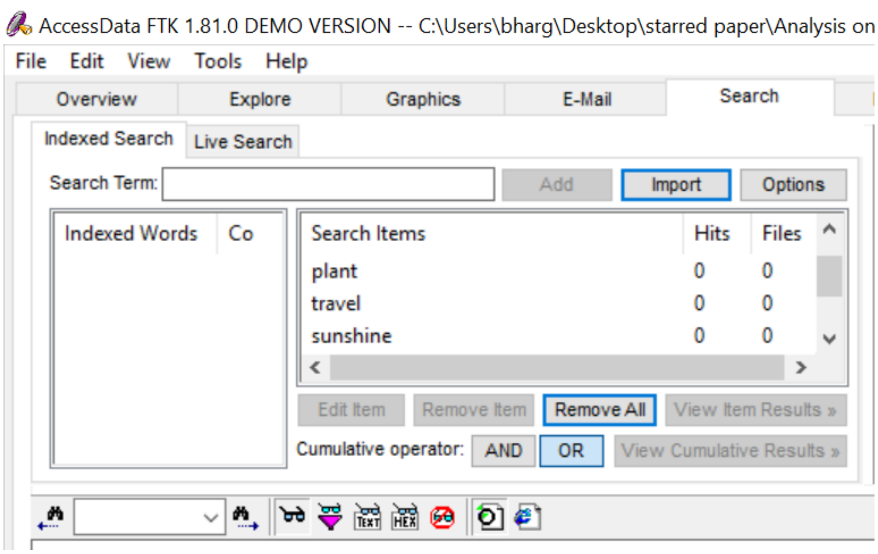

Figure 76. Sample Hits for Sunshine, Travel, and Plant for SSD Image 4

Now we search for .doc, .pdf and .xlxs files, and we see the number of hits as 0, and we can see as follows.

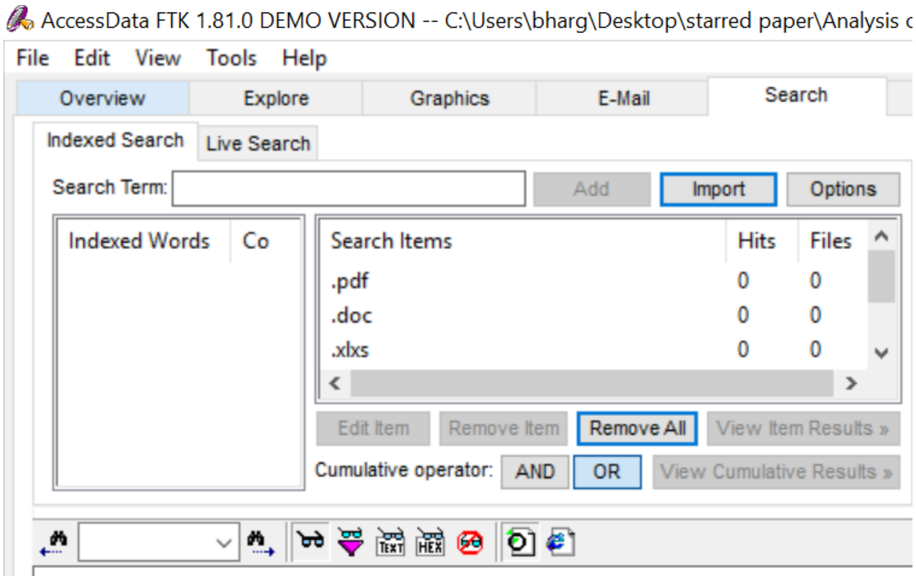

Figure 77. Sample Hits for .Doc, .Pdf, and .Xlxs for SSD Image 4

Now when we search for a specific item (animal), we get the hits as 0 which is far less than the results obtained from previous images.

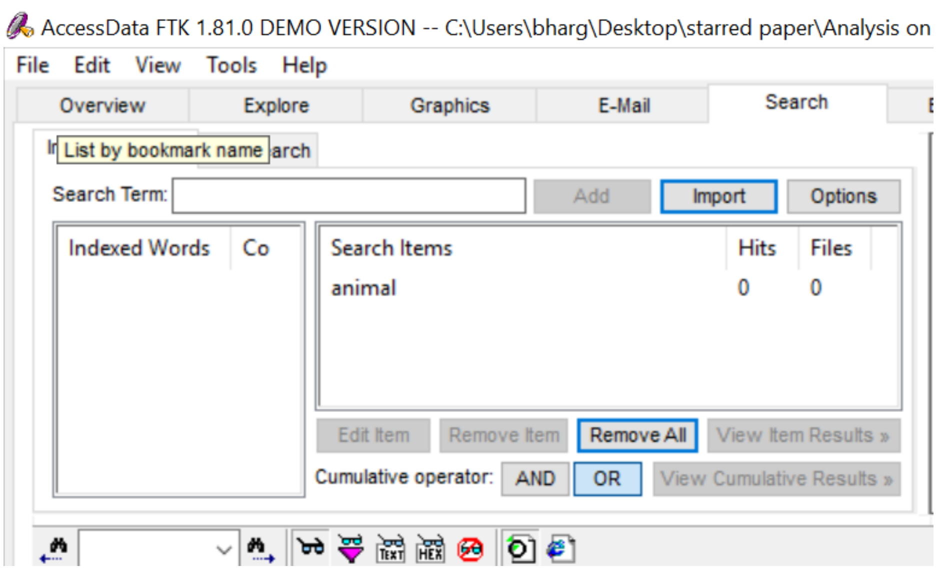

Figure 78. Sample Hits for Animals, Travel, and Plant for SSD Image 4

Results obtained from HDD. Now in the Search Term box, we search for indexed Words. From the below image we can see a sample number of hits for the search items animals, travel, plant, and sunshine.

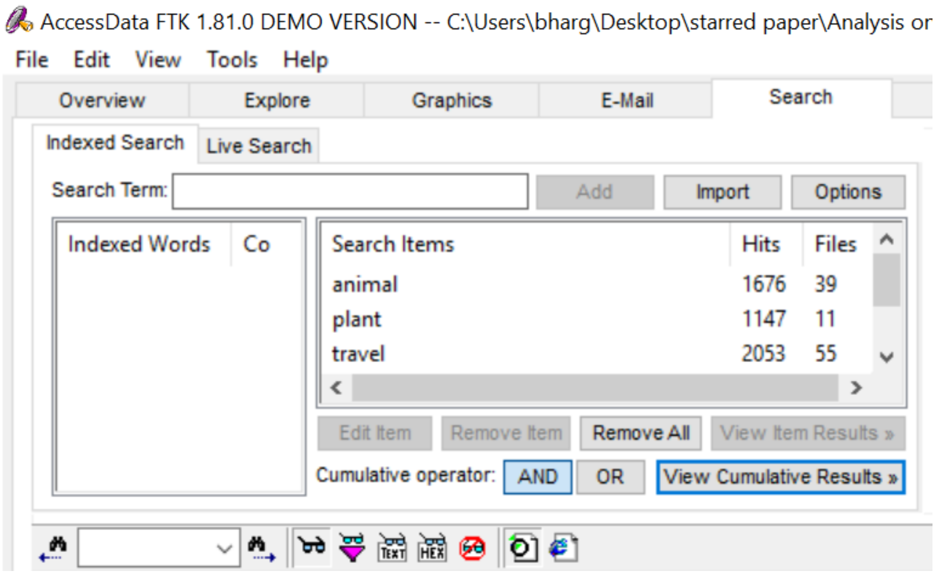

Figure 79. Sample Hits for Animals, Travel, and Plant for HDD Image

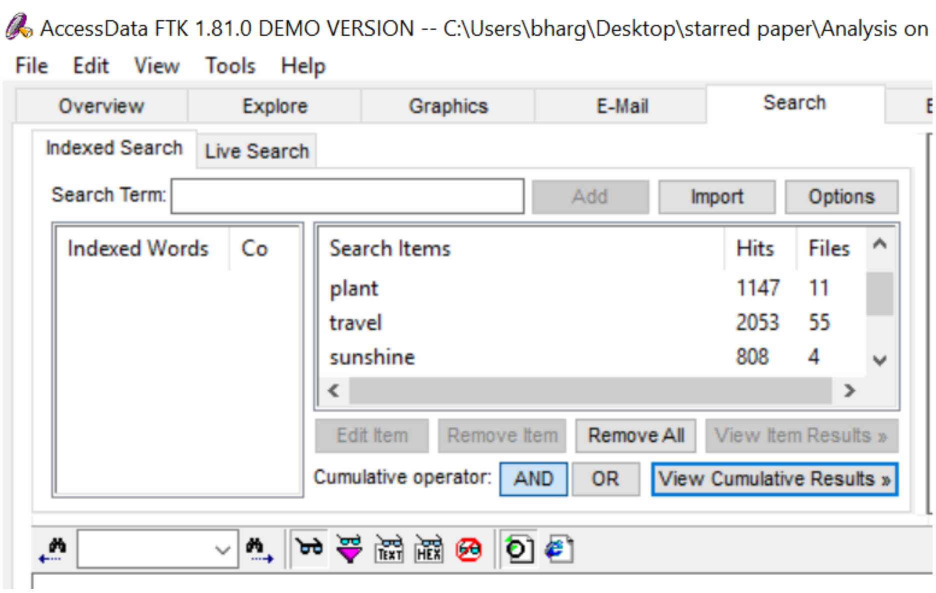

Figure 80. Sample Hits for Sunshine, Travel, and Plant for HDD Image

Now we search for .doc, .pdf and .xlxs files, and we see the number of hits as 0, and we can see as follows.

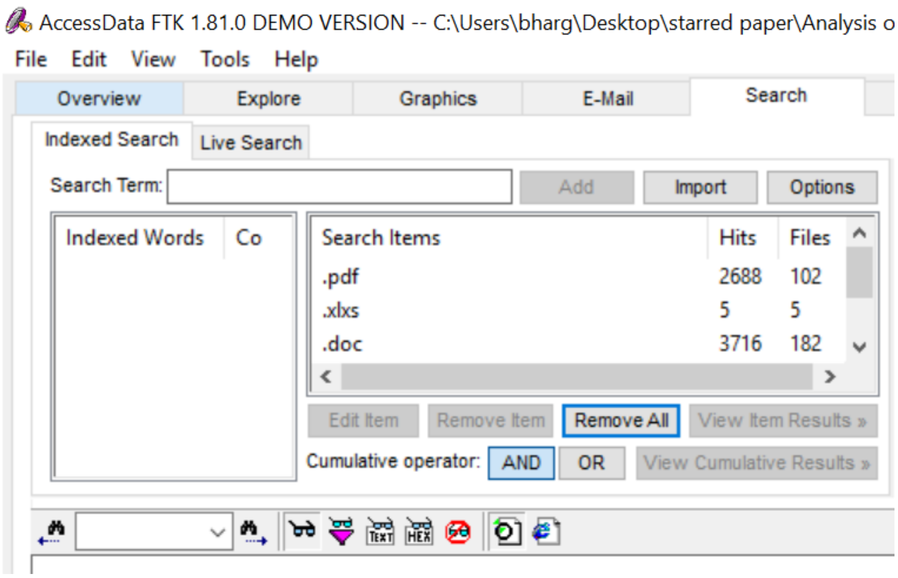

Figure 81. Sample Hits for .Doc, .Pdf, and .Xlxs for HDD Image

When we analyze the Hard Disk Drive using FTK, we can see the total number of files as of 5000.

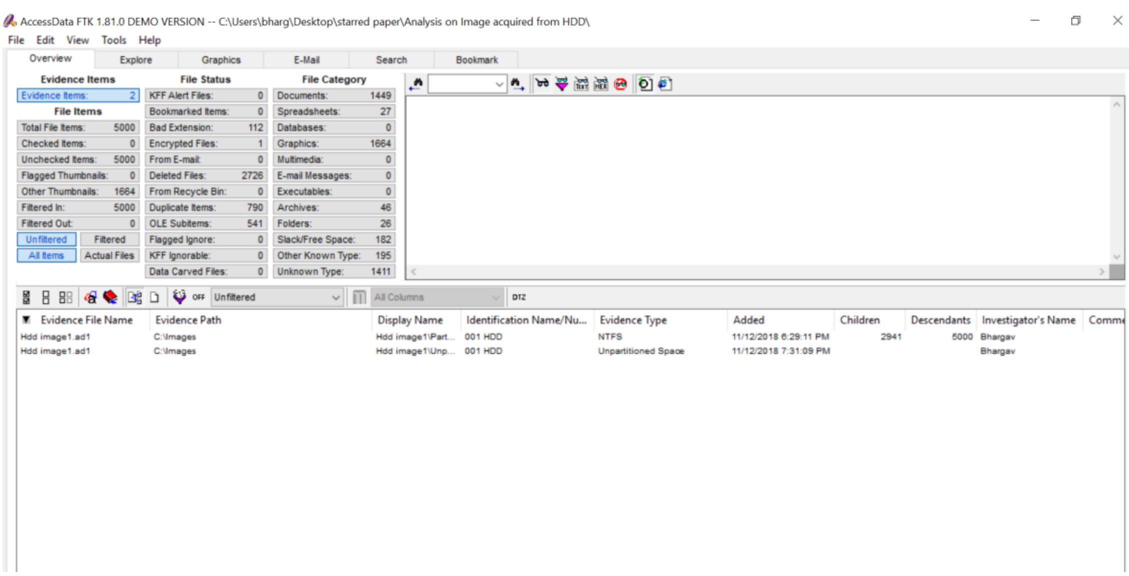

Figure 82. Total Number of Files Generated for HDD Image

When we click on one of the image files, we can recover it, and we can see the preview of the images as follows.

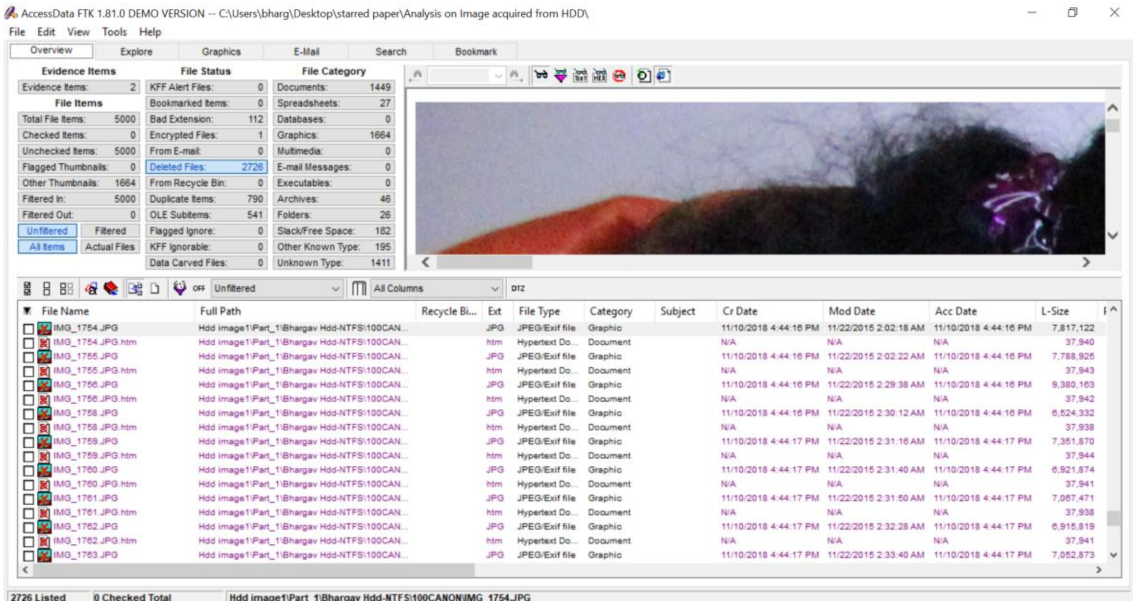

Figure 83. Image Preview for HDD Image

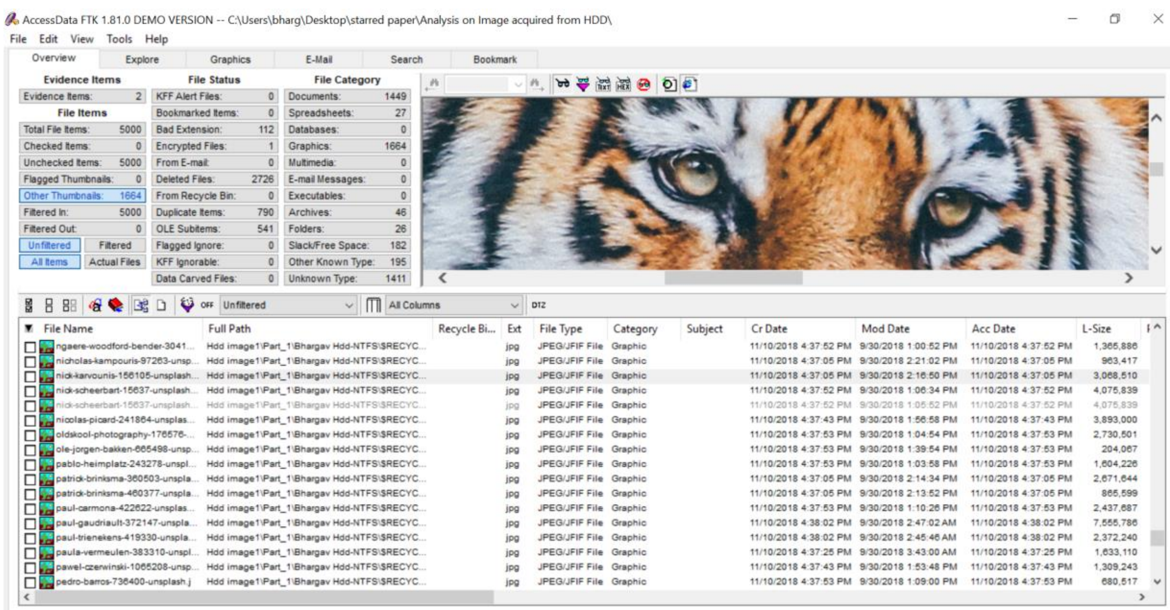

Figure 84. Image Preview 2 for HDD Image

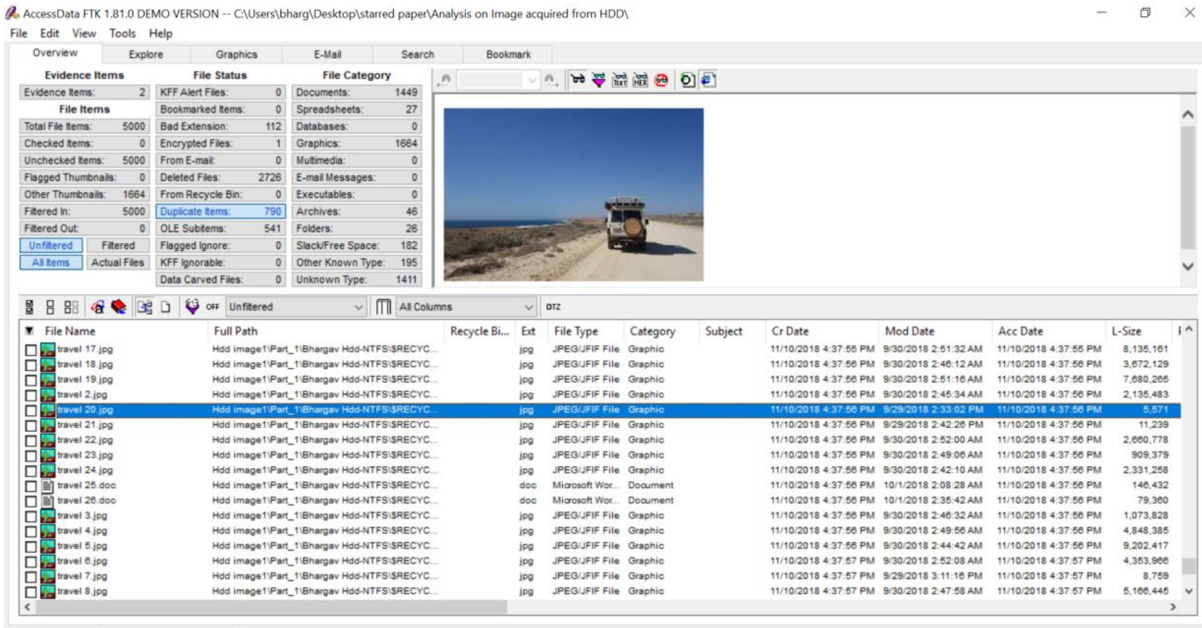

Figure 85. Image Preview 3 for HDD Image

When we click on one of the excel files, we can recover it, and we can see the preview of the excel as follows.

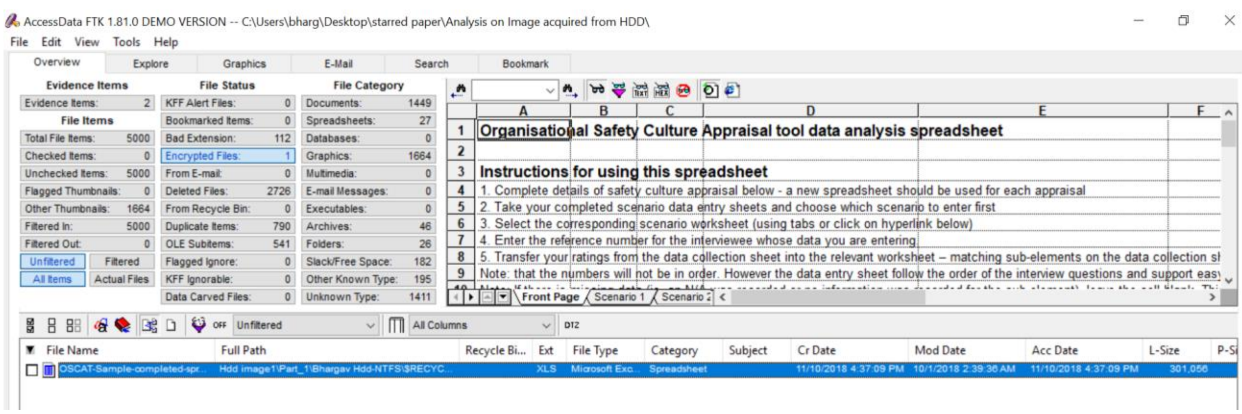

Figure 86. Excel Preview for HDD Image

When we click on one of the document files, we can recover it, and we can see the preview of the document as follows.

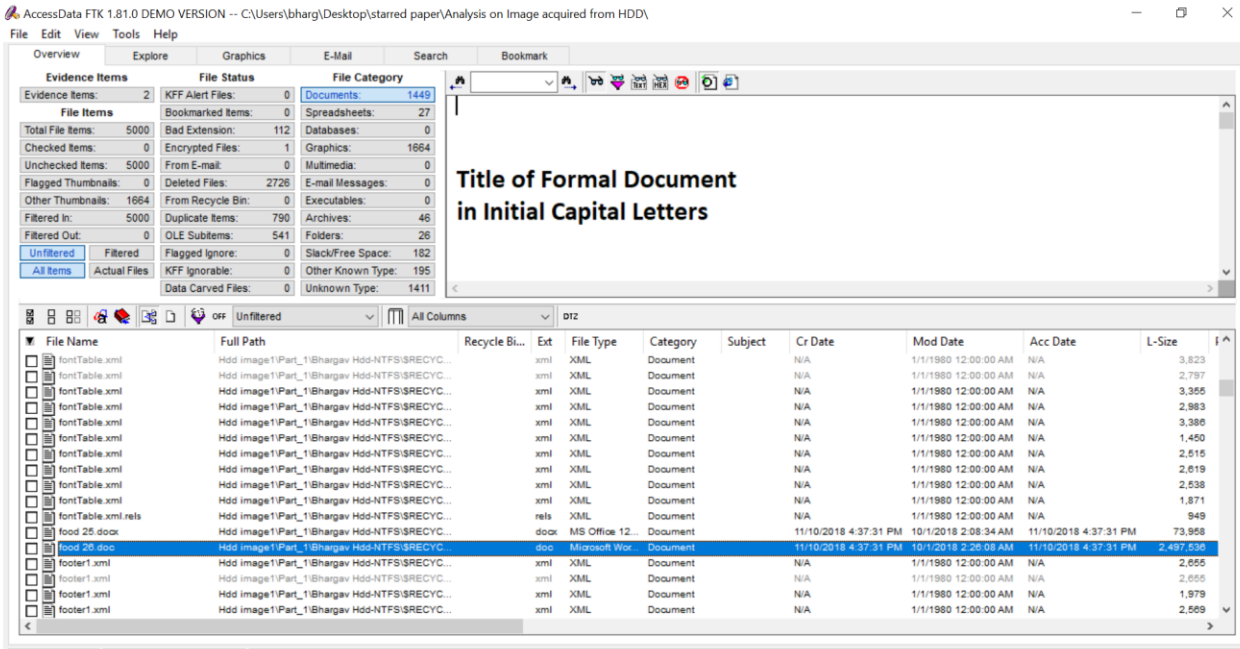

Figure 87. Document Preview for HDD Image

When we click on one of the folders, we can recover it, and we can see the preview of the folder in Hex code format as follows.

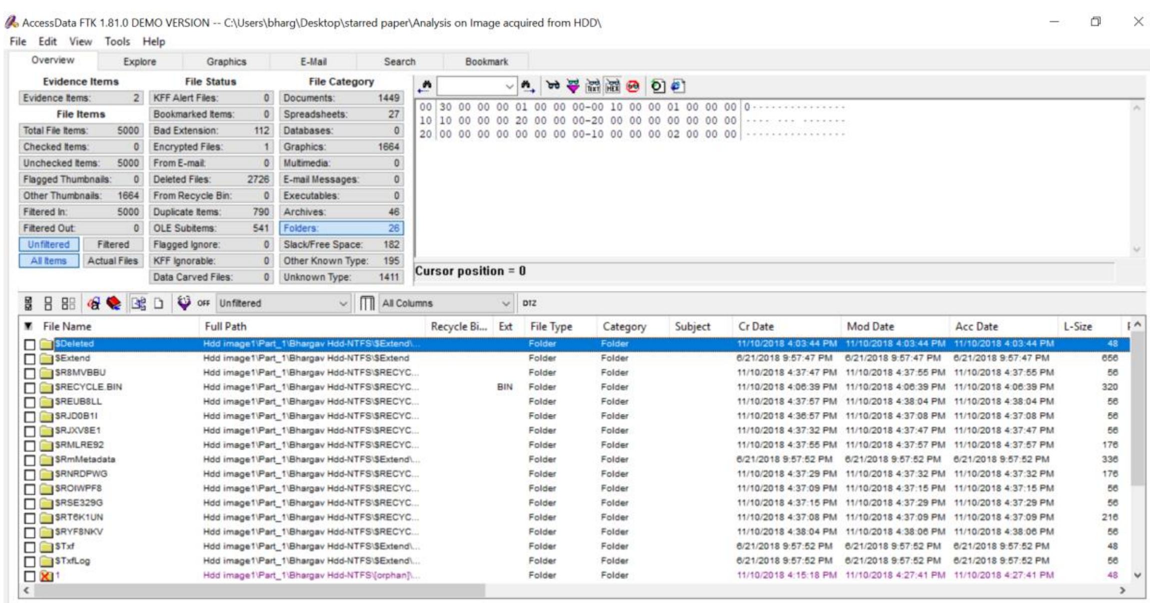

Figure 88. Folders Hex Code for HDD Image

Summary

In this chapter, we have loaded data into HDD and SSD and created images using FTK Imager. Next step is to load those images and analyze them as per digital forensic techniques. Analyze the results and conclusions are given in the next chapter.

Chapter V: Results, Conclusion, and Recommendations

Introduction

In this chapter, we will be comparing results and deriving the conclusion about the findings.

Results

Comparison of file hits in HDD vs. SSD.

Table 2

Comparison of Hits in HDD and SSD

	HDD 7/4/2018		SSD 7/4/2018	
Contents	No of Hits	No of files	No of Hits	No of files
Animal	1676	39	249	42
Travel	1147	11	763	95
Plant	2053	55	426	33
Sunshine	808	4	37	14
Doc files	3716	182	584	1498
PDF files	2688	102	161	1726
Excel files	5	5	399	366

Comparison of file hits in various images of Solid-State Drives.

Table 3

Comparison of File Hits in Various Images of Solid-State Drives

	7/4/2018 – Day 1 SSD1		7/14/2018 - Day 10 - SSD 2		7/24/2018 - Day 20 - SSD 3		8/4/2018 - Day 30- SSD 4	
Contents	No. of Hits	No. of Files	No. of Hits	No. of Files	No. of Hits	No. of Files	No. of Hits	No. of files
Animals	249	42	23	5	12	6	0	0
Travel	763	95	2	1	1	1	0	0
Plant	426	33	5	4	0	0	0	0
Sunshine	37	14	9	6	6	4	0	0
Doc files	584	1498	1420	1206	1245	1073	0	0
PDF files	161	1726	235	228	1971	1310	0	0
Excel files	399	366	4519	1441	260	245	0	0

Pictorial representation.

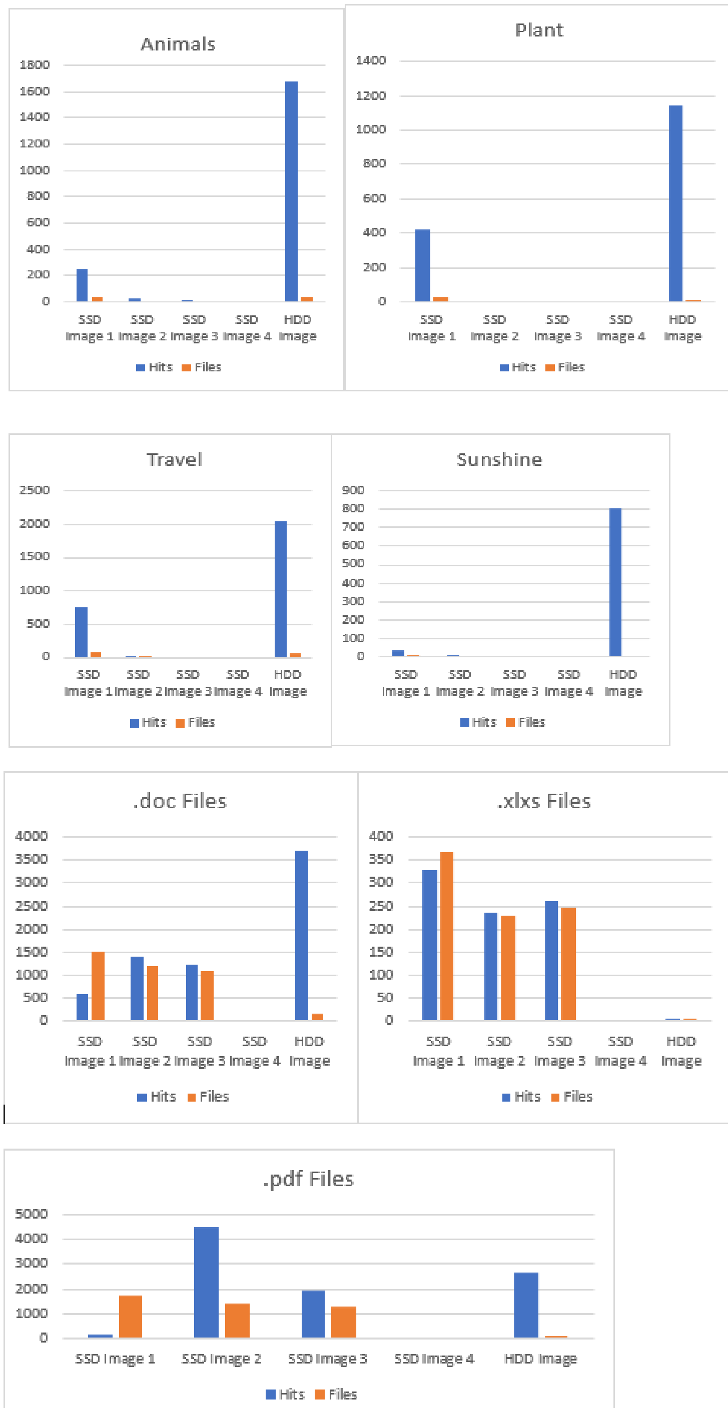

Figure 89. Comparison of All Images vs. Recovered Data

Conclusion

Based on our findings we clearly understand that the performance of SSD's is far beyond expectations to that of HDD's. When we use FTK to analyze our results, we can see that we can recover deleted files from the HDD. However, in the case of SSD, we have created four images, and we can see that the file count has been decreasing over time. It becomes a massive problem to the investigators as they cannot recover the deleted data.

When the forensic investigators collect the SSD as evidence, it is not possible for them to start the investigation right away. It has many processes before they can start looking at the evidence. By this time the data in the SSD will have been vanished and would not have any critical pieces of evidence to prove the case. There need to be some advanced tools and techniques to retrieve data from these devices.

Future Work

As a future work, it is recommended to come up with a change in the design of the solid-state drive with a switch which can disable the pre-clearing function of the drive. With the use of this switch, it becomes easier to stop automatic erasing of the data from the drive and preserve the data for the forensic experts. The switch can be implemented in the form of a magnetic switch integrated into the solid-state drive package which can be activated via a magnet that is placed outside this package case. When the switch is activated, this will give a signal to the processor or the controller to disable the pre-clearing function of the solid-state drive.

References

Bednar, P., & Katos, V. (2011). SSD: New challenges for digital forensics. In A. D'Stri, D. Te'enie, & M. De Marco (Eds.), *Information systems: A crossroads for organization, management, accounting, and engineering*. Retrieved from https://portal.research.lu.se/ws/files/5456453/4318024.pdf

Bell, G. B. (2010). Solid state drives: The beginning of the end for current practice in digital forensic recovery. *The Journal of Digital Forensics, Security and Law*, *5*(3), 1-20.

Brendan, H. (2017). *The battle between SSD and HDD is over, and the winner is clear.* Retrieved from https://www.yahoo.com/tech/battle-between-ssd-hdd-over-141508916

Bux, W., & Iliadis, I. (2010). Performance of greedy garbage collection in flash-based solid-state drives. *Performance Evaluation, 67*(11), 1172-1186.

C/net. (n.d.). *Digital storage basics, Part 4: SSD explained*. Retrieved from Cnet: https://www.cnet.com/how-to/digital-storage-basics-part-4-ssd-explained/

Chen, F. K. (2011). Making the best use of solid state drives in high-performance storage systems. In Proceedings of the International Conference on supercomputing, (ACM), pp. 22-32.

Geier, F. (2015). *The differences between SSD and HDD technology regarding forensic investigations*. Sweden: Linnaeus University. Retrieved from http://www.gti.bh/Library/assets/fulltext01-gshhsy652.pdf

INFOSEC. (n.d.). 7 best computer forensics tools. Retrieved from http://resources.infosecinstitute.com/7-best-computer-forensics-tools/#gref

King, C., & Vidas, T. (2011). Empirical analysis of solid state disk data retention when used with contemporary operating systems. *Digital Investigation, 8*(Supplement), S111-S117.

Phelps, J. (2012). How to partition and format your hard drive in windows. Retrieved from *PC World*, https://www.pcworld.com/article/248980/how-to-partition-and-format-your-hard-drive-in-windows.html

Ries, D. G., & Hill, C. (2017). Digical forensics in the courts. Retrieved from https://www.duq.edu/assets/Documents/forensics/_pdf/Forensic%20Fridays/March%2024%202017/Ries%20Digital%20Forensics%20Handout%203-24-17.pdf

Technopedia. (n.d.). In *Wikipedia*. Retrieved from https.//www.techopedia.com/

Thierolf, T., & Uriarte, J. (2010). Solid state drive architecture: A comparison and evaluation of data storage mediums. Retrieved from http://meseec.ce.rit.edu/551-projects/fall2010/1-4.pdf

Wei, M. G., Grupp, L. M., Spada, F. E., & Swanson, S. (2011). Reliably erasing data from flash-based solid state drives. In *Proceedings of 9th USENIX Conference on File and Storage Technologies (FAST '11)*, San Jose, CA.

Woodford, C. (2018). *Hard drives*. Retrieved from https://www.explainthatstuff.com/harddrive.html

Zhao, K. Z., Zhao, W., Sun, H., Zhang, T., Zhang, X., & Zheng, N. (2013). LDPC-in-SSD: Makng advanced error correction codes work effectively in solid state drives. In *Proceedings of 11th USENIX Conference on File and Storage Technologies* (FAST'13), San Jose, CA.