Data Security and Privacy

DSE 3258

L11 –Acquisition and Duplication

What is Computer Forensics?

- Computer Forensics is the science of obtaining, preserving, and documenting evidence from digital electronic storage devices, such as computers, digital cameras, mobile phones, and various memory storage devices.
- Digital evidence is any information or data of value to an investigation that is stored on, received by, or transmitted by an electronic device.
- **Text messages, emails, pictures and videos, and internet searches** are some of the most common types of digital evidence.
- **Computer forensics** is a set of **methodological methods** and **techniques** for identifying, gathering, preserving, extracting, interpreting, documenting, and presenting evidence from computing equipment in legal or administrative proceedings
- Computer Forensics is primarily concerned with the proper acquisition, preservation and analysis of digital evidence.

Acquisition and Duplication

- Data acquisition is the process of gathering evidence or information.
- This can be done by using established methods to acquire data from a suspected storage media outlet to gain access to information about the crime or other incident, and potentially using that data as evidence to convict a suspect.
- It is the use of established methods to extract Electronically Stored Information (ESI) from suspect computer or storage media to gain insight into a crime or an incident.
- It is a critical step in digital forensics, as an improper acquisition may alter data in the evidence media, and render it inadmissible in a court of law.
- Investigators must be able to verify the accuracy of acquired data, and the complete process should be auditable and acceptable in the court.

Acquisition and Duplication

- **Types of Data Acquisition**

➢ Live Acquisition

- Involves collecting data from a system that is powered ON

➢ Dead Acquisition (Static Acquisition)

- Involves collecting data from a system that is powered OFF

https://johntai.net/posts/chfi-notes/module-04/

Live Acquisition

- Live data acquisition involves collecting volatile data from a live system.
- Volatile information assists in determining the logical timeline of the security incident, and the possible users responsible.
- Live acquisition can then be followed by static/dead acquisition, where an investigator shuts down the suspect machine, removes the hard disk, and then acquires its forensic image.
- Types of data captured during live acquisition

➢ System Data

- Current configuration, Running state, Date and time, Command history Current system uptime, Running processes, etc

➢ Network Data

- Routing tables, , Network configuration, Network connections, etc.

- Live acquisition can help investigators obtain : Data from unencrypted containers or disks that are open on the system, automatically get encrypted when the system shuts down
- Private browsing history and data from remote storage services such as Dropbox (cloud service) by examining the Random-Access Memory (RAM)

Dead Acquisition

- Dead acquisition is defined as the acquisition of data from a suspect machine that is powered off
- Dead acquisition usually involves acquiring data from storage devices such as hard drives, DVD-ROMs, USB drives, flashcards, and smartphones
- Examples of static data: emails, word documents, web activity, spreadsheets, slack space, unallocated drive space, and various deleted files

Rules of Thumb for Data Acquisition

- **Do not work on original digital evidence**. Create a bit-stream/logical image of a suspicious drive/file to work on.
- Produce two or more copies of the original media
  - The first is the **working copy** to be used for analysis
  - The other copies act as the **library/control copies** that are stored for **disclosure** purposes or in the event that the working copy gets corrupt
- Use **clean media** to store the copies
- Upon creating copies of original media, verify the **integrity of copies** with the original

Types of Data Acquisition

➢**Logical Acquisition and Sparse Acquisition**

- In a situation with time constraints and when the investigator is aware of what files need to be acquired, logical acquisition is an ideal method.
- Logical acquisition allows an investigator to capture only selected **files** or **files types** of interest to the case
- Examples of logical acquisition include:
  - Email investigation that requires collection of **Outlook .pst or .ost files**
- Sparse acquisition is similar to logical acquisition, which in addition **collects fragments of unallocated data, allowing investigators to acquire deleted files**. Use this method when inspection of the entire drive is not required.

Types of Data Acquisition

➢**Bit-Stream Image**

- Bit-stream imaging creates a **bit-by-bit** copy of a suspect drive, which is a cloned copy of the entire drive including all its sectors and clusters.
- This image contains not just a copy of all the files and folders, but also the ambient data, which allows forensic investigators to **retrieve deleted files or folders.**

➢**Bit-stream disk-to-image file**

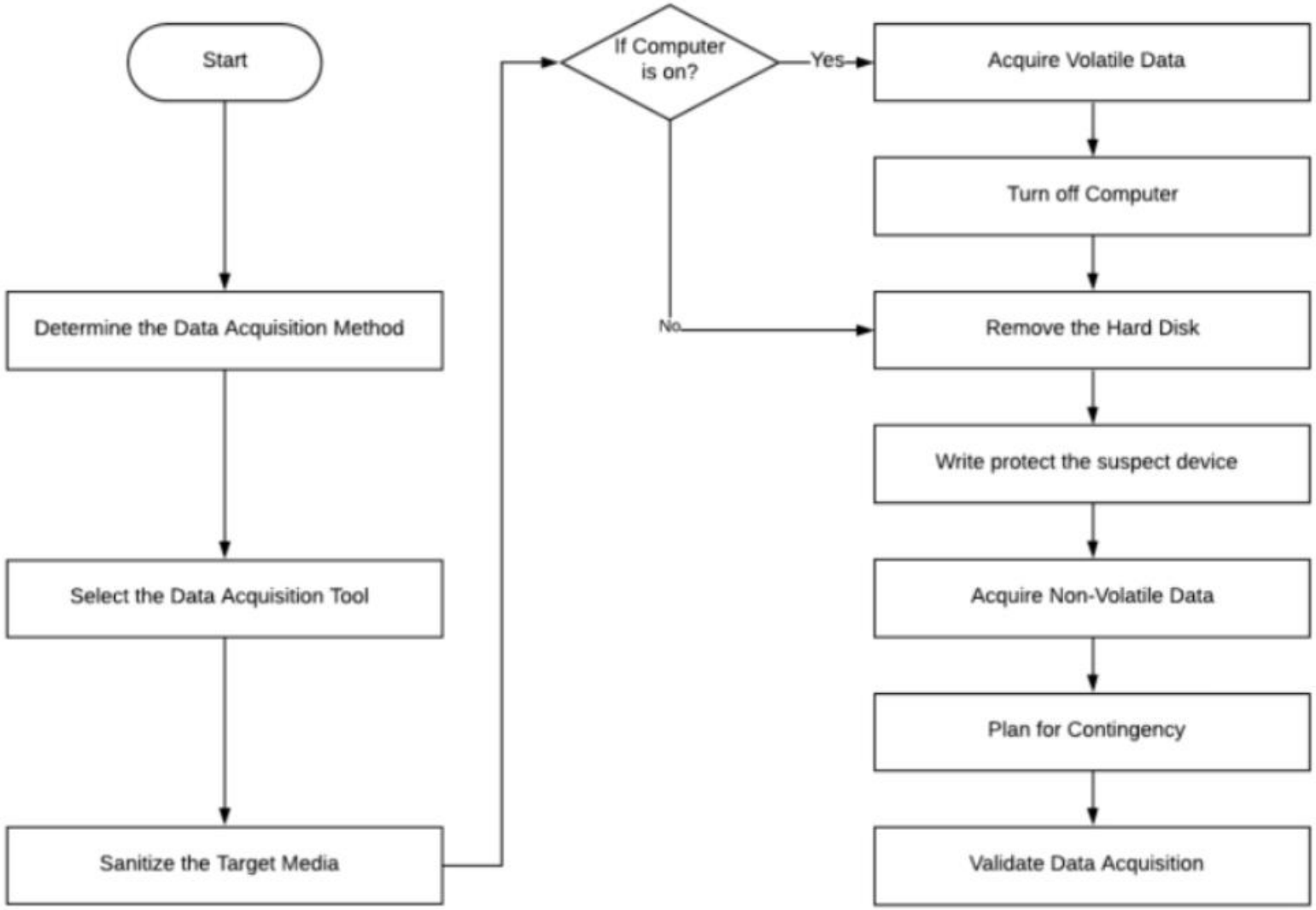
- It is the most common method used by **forensic investigators**
- With this method, one or many copies of the suspect **drive** can be generated
- The created image file is a bit-by-bit replica of the suspect drive
- Tools used: ProDiscover, EnCase, FTK, etc.

Types of Data Acquisition

➢**Bit-stream disk-to-disk**

- Disk-to-image copying is not possible in situations where
  - The suspect drive is very old and incompatible with the imaging software
  - Investigator needs to recover credentials used for websites and user accounts
- To overcome this situation, investigators can create a **disk-to-disk bit-stream** copy of the target media
- Tools used for this imaging process such as **EnCase, tableau Forensic Imager**, etc. enable investigators to modify the internal components of the target disk so that data obtained aligns well with the suspect drive.

Data Acquisition Methodology



Data Acquisition Methodology

Step 1: Determine the Best Data Acquisition Method

- An investigator needs to identify the best data acquisition method suitable for the investigation, depending on the investigator's situation. These situations include
  - ✓ Size of the suspect's drive
  - ✓ Time required to acquire the image
  - ✓ Whether the investigator can retain the suspect's drive
- Investigators need to acquire only the data that is intended to be acquired

Data Acquisition Methodology

Step 2: Select the Data Acquisition Tool

- Investigators need to choose the right tool for data acquisition based on the type of acquisition technique they choose. When it comes to imaging tools, they need to choose the tools that satisfy certain requirements.
- **Mandatory Requirements**
  - ✓ The tool should not change the original content
  - ✓ The tool should log I/O errors in an accessible and readable form, including the type of the error and location of the error
  - ✓ The tool must have the ability to pass scientific and peer review. Results must be repeatable and verifiable by a third party if necessary.
  - ✓ The tool should alert the user if the source is larger than the destination
  - ✓ The tool should create a bit-stream copy of the original content when there are no errors in accessing the source media
  - ✓ Tool documentation should be correct, i.e., the user should get expected results by implementing it in accordance with the tool's documented procedures

Data Acquisition Methodology

Step 3: Sanitize the Target Media

- ✓ Investigators must properly sanitize the target media in order to delete any prior data residing on it, before it is used for collecting forensic data
- ✓Post investigation, they must dispose of this media by following the same standards, so as to mitigate the risk of unauthorized disclosure of information and ensure its confidentiality

Step 4: Acquire Volatile Data

- ✓Volatile data acquisition involves collecting data that is lost when the computer is shut down or restarted
- ✓This data usually corresponds to running processes, logged-on users, registries, open files, etc.
- ✓While most of this data is recovered by examining the live system, the approximately same amount of data can be obtained by examining the image acquired from the memory of the system.

Data Acquisition Methodology

Step 5: Enable Write Protection on the Evidence Media

- ✓It is necessary to write protect the suspect drive using write blockers to preserve and protect the evidence contained in it.
- ✓A write blocker is a hardware device or software application that allows data acquisition from the storage media without altering its contents.
- ✓It blocks write commands, thus allowing read-only access to the storage media

Step 6: Acquire Non-volatile Data

- ✓Non-volatile data can be acquired in both live acquisition and dead acquisition. It mainly involves acquiring data from a hard disk.

- ✓There is no significant difference in the amount of data acquired from a hard disk between the live and dead acquisition methods.

Data Acquisition Methodology

Step 7: Plan for Contingency

- Investigators must prepare for contingencies such as when the hardware or software does not work, or a failure occurs during acquisition
  - Hard Disk Data Acquisition
    - ✓ Investigators must create at least two images of the digital evidence collected, in order to preserve it. If one copy of the digital evidence recovered becomes corrupt, investigators can then use the other copy.
  - If you have access to two or more imaging tools, you must create two images of the evidence using at least two of them. In case, you have access to only one tool, make two or more images of the drive using the same tool.
  - Be prepared to deal with encrypted drives that need the user to provide the decryption key for decrypting. Microsoft includes a full disk encryption feature (BitLocker) with select editions of Windows Vista and later.

Data Acquisition Methodology

Step 8: Validate Data Acquisition

- ✓Validating data acquisition involves calculating the hash value of the target media and comparing it with its forensic counterpart to ensure that the data is completely acquired

- ✓Hash value calculation generates a unique numeric value for files which is used for preserving data integrity and preventing data alteration

- ✓If two files contain the same hash value, they are taken to be completely identical even if those are named differently

Sterilizing evidence media

- Any previous data must be removed from the copy media with a software tool that is proven to remove all data from the drive.
- If media is not sterilized properly, the forensic output would be contaminated and in case of disposal of media, sensitive data may be leaked.
- By using forensically sterile media, the Computer Forensics specialist ensures that the media itself will not contaminate the evidence.
- This sterilization process should be documented and visually verified by the forensic examiner, leaving no doubt that whatever is found on the working copy during a forensic examination is/was also present on the original media.

Acquiring Forensics Images

- A forensic image (forensic copy) is a bit-by-bit, sector-by-sector direct copy of a physical storage device, including all files, folders and unallocated, free and slack space.
- Forensic images include all the files visible to the operating system (OS), as well as deleted files and pieces of files left in the slack and free space.
- This includes both the logical file structure (files and folders) and all the associated metadata for that logical structure.
- Electronic evidence can be gathered from a variety of sources, including computers, mobile devices, remote storage devices, internet of things (IoT) devices, and virtually any other computerized system.
- This image is created using various third-party tools which can easily capture the image of a hard drive bit by bit without changing even a shred of data.
- Forensic software copies data by creating a bitstream which is an exact duplicate.
- The best thing about creating a forensic image is that it also copies the deleted data, including files that are left behind in swap and free spaces.

Acquiring Live Volatile Data

- Volatile memory or random access memory stores information such as running process, incognito browsing sessions, clipboard data , information stored in plain text files and much more.
- Volatile data refers to the information stored in a system's temporary storage areas, like the RAM or physical memory, and in active processes or services. This data is characterized by its transient nature; it exists only as long as the system is powered on and can be lost or altered upon shutdown or restart.
- Volatile memory contains the following system artifacts which gets lost when the device is restarted or shut down. The following artifacts can be expected out of the volatile memory acquisition process:
    1. System Process
    2. Running services
    3. Clipboard Information
    4. Browsing Sessions (Incognito Sessions)
    5. Passwords
    6. Accessed Files and Multimedia
    7. Chats/Running Application stored data
- During an incident response, an investigator has to analyze the suspected machines and the profitability of capturing live RAM. Device has to be unlocked authentically and RAM dumping modules has to be loaded according to the host operating system.

Acquiring Live Volatile Data

- Highly Volatile
  - **These types of evidence are unlikely to be recoverable after the system is powered down**. Examples include physical memory, running processes, running services, screen shots, active network sessions, operational drivers, system information, and mounted encrypted volumes.
  - Physical Memory: Provides the opportunity to examine and carve potential passwords, recent messages, partial documents, malicious processes, web history, financial data, phone numbers, contact information, etc.
  - Running Processes: Provides the investigator or auditor with a record of the processes that were running on the target computer at the time of the acquisition. This information can provide clues about what the suspect or victim was doing most recently.
  - Running Services: Furnishes insight into the system services that were running or stopped. For example - was the antivirus active, was the firewall running, was there a VPN in operation?
  - Screenshots: Gives information about the most recent user activity, images, videos, messages, documents, and open web pages.
  - Active Network Sessions: Affords insight into the connections to inside or outside services. These could be NAS devices, cloud infrastructures, accomplices, or compromised services.
  - Operational Drivers: Provides detailed information about which peripherals have been connected to the system. For example, cameras, GPS devices, USB devices, etc. that could be valuable to the investigation or audit.
  - System Information: Provides IP and MAC addresses, general system information to link this computer/device to the acquired evidence.
  - Mounted Encrypted Volumes: Access to information that may be vital for the investigation, yet only available while file systems are mounted and unlocked.

Acquiring Live Volatile Data

- Moderately Volatile
  - **These types of evidence might be recoverable, but the process can be complex, slow, or less accurate**. Examples include user events (login, shutdown), security events, registry events, recent images, recent multimedia, recent documents, actively inserted devices, recently inserted devices, web history, and email history.
  - User Events (Login, Shutdown): Provides information about when the systems were used, when users logged in and logged out. Gives investigators evidence that could be used in questioning users.
  - Security Events: Provides auditors and compliance officers with information about possible security violations, unsuccessful login attempts, and changes to important security settings that could affect operations.
  - Registry Entries: Delivers a wealth of information about Windows systems, security settings, application settings, and even user activities.
  - Recent Images, Recent Multimedia, and Recent Documents: Offers a glimpse at the most recent images, multimedia, and documents viewed and modified by users.
  - Actively Inserted Devices and Recently Inserted Devices: Provides quick access to information about inserted USB and other memory devices.
  - Web History: Gives investigators insight into the most recent browsing habits of users.
  - Email History: Gives investigators access to email history and address books in use by users.

Acquiring Live Volatile Data

- Possibly Volatile or Time-Sensitive
  - **These types of evidence are most likely recoverable using postmortem procedures, but the recovery may be delayed**. Examples include files and documents, drive images, directory structure, and installed applications.
  - Files and Documents: Certain files by type or content may provide immediate evidence to investigators or auditors. These files may have vital data related to the investigation or contain company proprietary data.
  - Drive Images: In some cases, the direct image of a logical volume may be essential either to preserve evidence or acquire evidence that may be lost during shutdown or only be available in a live environment.
  - Directory Structure: Taking a snapshot of a directory structure may provide information about user activities and tendencies.
  - Installed Applications: Can provide a glimpse into the tendencies and sophistication of the user.

Acquiring Live Volatile Data

- **The steps for acquisition are as follows:**
  1. Determine the state of the machine
  2. Identify the operating system
  3. Check for authentic device access
  4. Insert acquisition media
  5. Perform Volatile Memory Dump
  6. Collect SWAP, PAGEFILE.sys and system protected files
  7. Hash and verify the acquired files
  8. Create Investigator copies
- **Memory Acquisition Procedure**
  - It is host operating system dependent.

Acquiring volatile memory from windows OS

Pre- Acquisition Process

To acquire volatile memory of windows OS based system, investigator needs to ensure the following measures:

1.Host machine should not be connected to any external network

- Disconnect LAN/WIFI/Bluetooth connections by putting the device in airplane mode
- Ensure the device is unlocked without installing any password bypassing module

2. Remove any external connected devices

3. Host machine should be connected to a stable power source

https://medium.com/@lucideus/windows-volatile-memory-acquisition-forensics-2018-lucideus-forensics-3f297d0e5bfd

- Windows memory management stores volatile memory in multiple ways, an investigator needs to acquire the following volatile information for proper volatile memory acquisition:
  - ✓**Pagefile.sys:** is a paging file which windows uses to store virtual memory contents.
  - ✓**Hiberfil.sys:** is created when windows hibernation is enabled. It stores volatile memory contents when the system needs to enter or has entered hibernation.
  - ✓**Swapfile.sys**: is used to store the idle and non active process data .
- **Acquisition Process**
  - To perform acquisition dump the memory contents to file along with other on-disk virtual memory storage files.
  - Example tools: FTK Imager lite, dumpit ,Lime
  - Ensure the following:
    - ✓Acquisition module should be executed from an external device
    - ✓The memory dump should be stored in an external drive
    - ✓Always hash the acquired data for integrity
    - ✓Ensure external drive has adequate free space calculating pagefile, hiberfil, swapfile and memory dump.

Metadata extraction

What is meta data?

- Often referred to as data that describes other data, metadata is structured reference data that helps to sort and identify attributes of the information it describes
- Meta is a prefix that -- in most information technology usages -- means "an underlying definition or description." Metadata summarizes basic information about data, which can make it easier to find, use and reuse particular instances of data.

Descriptive Meta data

- Descriptive metadata is basic information: who, what, when and where.
- Think of it as a description of a file or a piece of art with the plaque next to it; it's there to help individuals know what they are looking at and the description changes depending on the contents of the object or information piece.
- Types of descriptive metadata include:
  - Time and date of creation
  - Program or processes used for the creation of the data
  - Purpose of the data
  - Creator or author of the data
  - Location on a device where the data was created
  - Technical standards used
  - File size
  - Data quality
  - Source of the data
  - Modifications or programs used to modify the file

Metadata extraction

There are different types of metadata that can be found in a digital file, these include:

- **File Name:** This is the name given to the file when it was originally created and can be a useful way of identifying its source. This can also help identify a specific version of malware if there are multiple iterations or variants available in an attacker's toolkit.
- **File Size:** This can be used to identify the size of a file and whether it has been modified or tampered with. If a file is modified by an attacker, then there will be some changes in its size compared to the original. This can also be useful for identifying new versions of malware if they are released by their authors.
- **Date Modified:** The date a file was last modified can be useful in helping to identify if it has been modified by an attacker. If a file is not modified, then there will be no change in its date of modification.
- **Location on Disk:** If a file is stored in a compressed form, it can be identified by its location on disk. For example, if the file is stored in the same folder as other files that were not modified by an attacker and it has been modified, then it would be likely that this file was modified by someone else.
- **File Hash:** If a file has been modified by an attacker, it will have a different hash value than the original. A hash value is a unique number that can be used to identify particular files. This can be helpful in identifying if a file has been modified by an attacker.

Why Meta data is useful in cyber forensic?

- Metadata examination is extremely useful in the field of cyber forensics, especially if the metadata contains information that is not easily obtainable. When a file is moved from one directory to another, the modification time and access time may change, however the creation time will remain the same (if the OS supports it). The hash value of a file can be used to determine if a file has changed since its inception. If there are no changes made to a file from the time it was created until now then these values should be identical.
- For example, say that you have a spreadsheet containing data from your company's sales for the third quarter of 2022. The metadata on this file will tell you when it was created as well as any changes made since then.
- Metadata can be very useful in cyber forensics because it provides investigators with more information than simply retrieving the contents of files or computer hardware devices. Metadata can help investigators determine if files have been altered since they were first created or if they were written at all – something that may not be easily found through other means such as timestamping or hashing algorithms

Metadata extraction

- More sophisticated metadata might also contain data summaries or keywords, content ratings, location coordinates, or system file identification labels.
- Verifying metadata is crucial in digital forensic investigations as it helps establish the authenticity, integrity, and reliability of the evidence.
- Metadata extraction is a tool for understanding the content, context, and trends within large volumes of data—data that can expose insights about potentially malicious threats.

File-system analysis

- File systems analysis is a fundamental aspect of digital forensics, involving the understanding of how data is stored, organized, and retrieved on storage media such as hard drives, solid-state drives, and removable storage devices. Here's an overview of file systems analysis and the key concepts involved:
- File System Basics:
  - Definition: A file system is a method used by operating systems to organize and store data on storage devices, providing a structured way to store, retrieve, and manage files.
  - Components: A file system consists of various components, including the boot sector, file allocation table (FAT), master file table (MFT), inode table, directory structure, and file metadata.

File-system analysis

- Storage Media Structures:
  - Boot sector: The boot sector is the first sector of a storage device and contains the boot loader program responsible for booting the operating system.
  - Partition table: The partition table stores information about the partitions on a storage device, including their size, location, and file system type.
  - File allocation table (FAT): FAT is a file system used by older versions of Windows to track the allocation of disk space to files and directories. It consists of entries that map file clusters to file names and attributes.
  - Master file table (MFT): The MFT is a key component of the NTFS file system used by modern versions of Windows. It contains metadata about files and directories, including file attributes, timestamps, and data run extents.
  - Inode table: Inode-based file systems like ext2/ext3/ext4 used in Linux store file metadata and pointers to data blocks in a data structure called an inode table.

File-system analysis

- File System Analysis Techniques:
  - Data structure parsing: Analysts parse and interpret the data structures of file systems to extract information about files, directories, and metadata.
  - File signature analysis: Analysts identify file types and formats by analyzing file signatures or magic numbers, which are unique identifiers found in file headers.
  - Unallocated space analysis: Analysts examine unallocated space on storage media to recover deleted or fragmented files and identify remnants of past user activity.
  - Timestamp analysis: Analysts analyze timestamps associated with files and directories (e.g., creation time, modification time, access time) to reconstruct timelines of user activity.

File-system analysis

- Directory Structure Analysis:
  - Directory hierarchy: Analysts analyze the hierarchical structure of directories to understand how files are organized and stored on a storage device.
  - Directory entry parsing: Analysts parse directory entries to extract information about file names, attributes, timestamps, and file paths.
- Metadata Analysis:
  - File metadata: Analysts examine file metadata such as file attributes (e.g., read-only, hidden, system), timestamps (e.g., creation time, modification time), and file size.
  - Directory metadata: Analysts analyze directory metadata such as timestamps, permissions, and file system quotas to understand directory properties and access controls

File-system analysis

- Forensic Significance:
  - Reconstruction of digital events: File systems analysis enables analysts to reconstruct digital events, such as file creation, modification, and deletion, to establish timelines of user activity and potential evidence tampering.
  - Identification of relevant artifacts: File systems analysis helps identify relevant artifacts and evidence for forensic investigations, including incriminating files, suspicious directories, and hidden data.